Computational Studies Applied to Model Systems for Proteins and Materials

Risnita Vicky Listyarini

Mestrado em Química

Departamento de Química e Bioquímica

2016

Orientadora

Maria João Ramos, Professora Catedrática, Faculdade de Ciências

Universidade do Porto

Coorientadores

Pedro A. Fernandes, Professor Associado, Faculdade de Ciências

Universidade do Porto

Natércia F.Brás, Investigadora, Faculdade da Ciências, Universidade do Porto

Ria Broer, Full Professor, Faculty of Mathematics and Natural Sciences, University of Groningen

Remco W. A. Havenith, Lecturer, Faculty of Mathematics and Natural Sciences, University of Groningen

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.
O Presidente do Júri,

Porto, _____/_____/_____

ACKNOWLEDGEMENTS

This research project would not have been possible without the support of many people. I would like to express my gratitude to all of the people who were helpful and offered me invaluable assistance support and guidance.

First, I would like to give special thanks to the Erasmus Mundus programme of Theoretical Chemistry and Computational Modelling (TCCM) with the support of the European Union. I am very thankful for the opportunity to receive the international European master degree and have the experience of doing the project in two countries.

I would like to express my gratitude for carrying out this project at the research group of Theoretical Chemistry, Department of Chemistry & Biochemistry, Faculty of Sciences, University of Porto. In particular, I would like to express my deepest gratitude to my supervisors Maria João Ramos, Pedro Alexandrino Fernandes and Natércia F. Bras.

I wish to express my appreciation also to another research group, for the opportunity of doing part of my research project at the Theoretical Chemistry Group of Zernike Institute for Advanced Materials, Faculty of Mathematics and Natural Sciences, University of Groningen. I am highly thankful to my supervisors Ria Broer, Remco W. A. Havenith and Riccardo Alessandri.

Finally, I wish to express my heartfelt thanks to my beloved family, supporting and encouraging me through the duration of my studies, my friends for their help and wishes for the successful completion of this project.

ABSTRACT

The current research project consists of two parts within the main framework of the computational studies of different systems. The first part, carried out at the group of Theoretical Chemistry of the University of Porto, is related to the benchmarking of density functionals for the electron affinity of the Fe(III)/Fe(II) redox pair. We have analysed the accuracy of Density Functional Theory for calculating the difference in electronic energy ($\Delta E'_{el}$) of the Fe(III)/Fe(II) pair, the most relevant contribution for the Fe(III)/Fe(II) redox potential. Forty-five density functionals were applied to sixteen model systems for Fe(III)/Fe(II), which are representative models of biological systems. Each structure of Fe(II) and Fe(III) was optimized at the MP2/6-311+G(d,p) level and the reference values for the benchmarking were determined at the very accurate CCSD(T)/CBS level. The total energy at the CBS limit was obtained as the sum of the complete basis set of MP2 energies and the CBS correlation energy calculated with the aug-cc-pVDZ basis set. The results from the DFT benchmarking led to the conclusion of which are the best functionals, within those evaluated, for a total of sixteen model systems.

The second part of the project, carried out at the group of Theoretical Chemistry of the University of Groningen, is related to the study of the impact of the polarity of fullerene derivatives on the morphology of bulk heterojunction (BHJ) solar cells from solvent evaporation simulations. The Martini Coarse-grained (CG) model combined with solvent evaporation simulations was used to simulate the fullerenes derivatives, PCBM, DMPCBM and DMPCHp. The main objective of the study was to test whether the simulations results agree with the experimental findings by Matsumoto *et al* [78]. Accordingly, the results of the currently performed studies showed an agreement with the experimental studies performed by Matsumoto *et al* [78]

Keywords: benchmarking DFT, electron affinity, electronic energy ($\Delta E'_{el}$), Fe(III)/Fe(II) redox pair, extrapolation complete basis set (CBS), bulk heterojunction (BHJ), fullerene derivatives, Martini Coarse-grain (CG) model, solvent evaporation simulations.

RESUMO

O presente projeto de investigação consistiu em duas partes relacionadas com estudos computacionais de sistemas diferentes. A primeira parte, realizada no grupo de Química Teórica da Universidade do Porto, relativa à avaliação comparativa de funcionais de densidade para o cálculo da afinidade eletrónica do par redox Fe(III)/Fe(II). Em particular, foi analisada a exatidão da teoria do funcional da densidade para calcular a diferença de energia eletrónica ($\Delta E'_{el}$) do par Fe(III)/Fe(II), basicamente a contribuição mais relevante para o seu potencial redox. Foram estudados quarenta e cinco funcionais de densidade e dezasseis sistemas modelo para o par redoxFe(III)/ Fe(II), representativos dos existentes em sistemas biológicos. Os valores de referência para o *benchmarking* foram determinados ao nível de exatidão muito elevado CCSD (T)/CBS. A energia total no limite CBS foi obtida como a soma da energia obtida com um conjunto completo de funções de base MP2 e a energia de correlação CBS calculada com um conjunto de funções de base aug-cc-pVDZ. Os resultados do *benchmarking* levaram à conclusão de quais os melhores funcionais, de entre os avaliados, para o estudo dos dezasseis sistemas modelo utilizados.

A segunda parte do projeto, realizada no grupo de Química Teórica da Universidade de Groningen, consistiu no estudo do impacto da polaridade de derivados do fulereno na morfologia do *heterojunction* (BHJ) em células solares através de simulações de dinâmica molecular da evaporação do solvente. O modelo *Martini Coarse-grain* (CG) e simulações de evaporação de solventes foram utilizados para estudar os derivados de fulerenos, PCBM, DMPCBM e DMPCHp. O principal objetivo do estudo foi o de verificar se os resultados das simulações concordavam com os valores experimentais obtidos por Matsumoto *et al* [78]. Os resultados dos estudos computacionais realizados demostraram estar de acordo com os estudos experimentais levados a cabo por Matsumoto *et al* [78]

Palavras-chave: DFT *benchmarking*, afinidade eletrónica, energia electrónica ($\Delta E'_{el}$), Fe(III)/Fe(II) par redox, extrapolação para o limite da base de funções (CBS), heterojunction granel (BHJ), derivados de fulereno, modelo Martini Coarse-grain (CG), simulações de evaporação de solventes.

TABLE OF CONTENTS

Acknowledgements i
Abstract.....ii
Resumo.....iii
Table of Contents iv
Index of Tables vi
Index of Figures vii
List of Abbreviations ix
I. Benchmarking of Density Functionals for the Electron Affinity of the Fe(III)/Fe(II) Redox Pair.....1
1. Introduction.....1
2. Literature Review.....5
2.1 Ionization potential (IP) and electron affinities (EA).....5
2.2 Computation Reduction Potential.....6
2.3 Theoretical Methods.....8
2.3.1 Moller-Plesset Perturbation Theory.....8
2.3.2 Single and Double Coupled Cluster Theory With Perturbative Triple Correction (CCSD(T)).....10
2.3.3 Extrapolation to Complete-Basis-Set.....11
2.3.4 Density Functional Theory (DFT)..... 13
3. Computational Details.....15
3.1 Model Systems.....15
3.2 Optimisation Structure.....17
3.3 Calculation of electronic energy E'_{el}.....17
3.4 Benchmarking of DFT functional.....18
4. Results and Discussion 20
4.1 Calculation of $\Delta E'_{el}$ at MP2/CBS.....20
4.2 Calculation of $\Delta E'_{el}$ at CCSD(T)/CBS level.....23
4.3 Benchmarking of DFT functional.....29
5. Conclusion.....30
6. Further Work.....39
II. Impact of Polarity of Fullerene Derivatives on the Morphology of Bulk Heterojunction Solar Cells from Solvent Evaporation Simulations.....40
1. Introduction.....40

2. Models and Methods....43
2.1 The ideal morphology for BHJ solar cells....43
2.2 Molecular Dynamics simulation....43
2.3 Coarse-grain Martini Force Field Model....44
3. Computational Details....46
3.1 Coarse-grain models for PCBM, DMPCBM and DMPCHp....46
1. PCBM....46
2. DMPCBM....48
3. DMPCHp....48
3.2 Simulation parameter....49
3.3 Simulations in solution....49
3.4 Solvent evaporation simulation....50
3.5 Morphological Analysis of fullerenes derivatives....51
4. Results and discussion....51
4.1 Coarse-grain models for PCBM, DMPCBM and DMPCHp....51
4.2 Simulations in solution....53
4.3 Solvent evaporation Simulation....56
4.4 Morphological Analysis of fullerene derivatives....57
4.5 Analysis number of contacts over time....61
5. Conclusion....66
6. Further Work....66
References....68
Supporting Information....77

Index of Tables

Table 1. List of DFT Functionals and basis sets used in this work……19
Table 2. $\Delta E'_{el}$ (Fe(II) - Fe(III)) at MP2 level using aug-cc-pVDZ, aug-cc-pVTZ and aug-cc-pVQZ basis set……...20
Table 3. $\Delta E'_{el}$ (Fe(II) - Fe(III)) at MP2/CBS using Truhlar Extrapolation Scheme…….21
Table 4. $\Delta E'_{el}$ (Fe(II) - Fe(III)) at MP2/CBS using the Helgaker Extrapolation Scheme and the difference between Truhlar and Helgaker Extrapolation Methods……...22
Table 5. $\Delta E'_{el}$ (Fe(II) - Fe(III)) at CCSD(T)/CBS, MP2/CBS and the difference between each method……..…….23
Table 6. Bond lengths and Mulliken atomic charges in Group A (Fe with water ligands).....25
Table 7. Bond lengths and Mulliken atomic charges in Group B (Fe-ligand X; X= CH_3O^- , $HCOO^-$, CH_3S^-, $C_6H_5O^-$)….....26
Table 8. $\Delta E'_{el}$ (Fe(II) - Fe(III)) CCSD(T)/CBS energy of of Group C…….....27
Table 9. Bond lengths in Group C ([Fe-$(H_2O)_3$(ligand X)]; (X= NH_2CH_3, H_2O, $HCOO^-$, OCH_3^- and CH_3S^-)……27
Table 10. $\Delta E'_{el}$ (Fe(II) - Fe(III)) CCSD(T)/CBS energy of Group D……...28
Table 11. Bond lengths in Group D ([Fe-$(H_2O)_5$-ligand X)]; (X= NH_2CH_3, H_2O, $HCOO^-$, CH_3O^- and CH_3S^-)……...29
Table 12. Ten density functionals with the best performance for Group A complexes……30
Table 13. Ten density functionals with the best performance for Group B complexes……32
Table 14. Ten density functionals with the best performance for Group C complexes…...33
Table 15. Ten density functionals with the best performance for Group D complexes…...34
Table 16. Ten density functionals with the best performance for all Fe-ligands systems...36
Table 17. The number of contacts in three different fullerene derivatives in ratio 0.7:1.....58
Table 18. The number of contacts P3HT-fullerene in three different fullerene derivatives in ratio 1:0.8……..60
Table 19. Standard deviation and standard error for number of contacts for ratio 0.7:1....60
Table 20. Standard deviation and standard error for number of contacts for ratio 1:0.8….61
Table 21. The number of contacts for ratio 0.7:1 of the 15x15x80 box……..62
Table 22. The number of contacts for ratio 1:0.8 of 15x15x80 box……63
Table 23. The number of contacts for ratio 1:0.8 of 20x20x80 box.....64
Table 24. The number of contacts for ratio 1:0.8 of 20x20x80 box (second batch).....65

Index of Figures

Fig. 1 - Thermodynamic cycle for the calculation of Gibbs free energies of one-electron reduction process.....2
Fig. 2 - Fig. 2 Thermodynamics cycle for the calculation of the absolute and relative reduction potentials for the Fe (III)/Fe(II) redox pair.....8
Fig. 3 - Jacob's Ladder of density functional approximation [25].....14
Fig 4. - Representation of the different complexes with Fe(II) and Fe(III), studied: (i) $Fe(H_2O)$ (ii) $Fe(H_2O)_2$ (iii) $Fe(H_2O)_4$ (iv) $Fe(H_2O)_6$ (v) $Fe(OCH_3)$ (vi) $Fe(SCH_3)$ (vii) Fe(COOH) (viii) $Fe(C_6H_5O)$ (ix) $Fe(H_2O)_3(OCH_3)$ (x) $Fe(H_2O)_3(SCH_3)$ (xi) $Fe(H_2O)_3(COOH)$ (xii) $Fe(H_2O)_3(NH_2CH_3)$ (xiii) $Fe(H_2O)_5(OCH_3)$ (xiv) $Fe(H_2O)_5(SCH_3)$ (xv) $Fe(H_2O)_5(COOH)$ (xvi) $Fe(H_2O)_5(NH_2CH_3)$.....16
Fig. 5 - MUE and MaxE values of the best 10 density functionals for systems of Group A.....31
Fig. 6 - MUE and MaxE values of the best 10 density functionals for systems of Group B.....33
Fig. 7 - MUE and MaxE values of the best 10 density functionals for systems of Group C.....34
Fig. 8 - MUE and MaxE values of the best 10 density functionals for systems of Group D.....35
Fig. 9 - MUE and MaxE values of the best 10 density functionals for all 16 systems.....36
Fig. 10 - Bulk Heterojunction (BHJ) solar cell. Picture was taken from de Gier *et al*. [88]..41
Fig. 11 - Processes which happen in a BHJ. Picture was taken from Wodo et al. [90].....41
Fig. 12 - (a) Typical morphology BHJ (b) ideal morphology BHJ. Picture was taken from Wodo *et al* [90].....43
Fig. 13 - Fullerene C_{60} (left) and PCBM structure (right).....46
Fig. 14 - PCBM, P3HT and chlorobenzene (CB) beads representation.....47
Fig. 15 - PCBM structure (left) and DMPCBM structure (right).....48
Fig. 16 - DMPCBM structure (left) and DMPCHp structure (right).....49
Fig. 17 - Solvation box for simulation inside solvent.....50
Fig. 18 - Solvent evaporation simulation was adapted from Alessandri [104].....51
Fig. 19 - Beads representation of PCBM (left) and DMPCBM (right).....52
Fig. 20 - Bond distance in centre of mass of 1,3-dimethoxybenzene (DMB).....53
Fig. 21 - Beads representation of DMPCBM (left) and DMPCHp (right).....53
Fig. 22 – The procedure of making CG trajectory from 21 beads into 2 beads.....54

Fig. 23 - Radial Distribution Function (RDF) of fullerene C_{60}……..….55
Fig. 24 - Radial Distribution Function (RDF) of fullerene C_{60} and sidechain…….56
Fig. 25 - Morphological evolution of PCBM in solvent evaporation simulation (red = P3HT, blue = fullerene)…..…….....56
Fig. 26 - Final morphological appearances of thin films DMPCHp, PCBM and DMPCBM using 0.7:1 (P3HT:fullerene) for 15x15x80 nm^3 box (red = P3HT, blue = fullerene). Area under yellow circle is the aggregation cluster of fullerene…….57
Fig. 27 - Final morphological appearances of thin films DMPCHp, PCBM and DMPCBM using 0.7:1 (P3HT:fullerene) for 20x20x80 nm^3 box (red = P3HT, blue = fullerene). Area under yellow circle is the aggregation cluster of fullerene……58
Fig. 28 - Final morphological appearances of thin films DMPCHp, PCBM and DMPCBM using 1:0.8 (P3HT:fullerene) for 15x15x80 nm^3 box (red = P3HT, blue = fullerene). Area under yellow circle is the aggregation cluster of fullerene.....59
Fig. 29 - Final morphological appearances of thin films DMPCHp, PCBM and DMPCBM using 1:0.8 (P3HT:fullerene) for 20x20x80 nm^3 box (red = P3HT, blue = fullerene). Area under yellow circle is the aggregation cluster of fullerene.....59
Fig. 30 - The percentage (%) P3HT-fullerene number of contacts for ratio 0.7:1 of 15x15x80 box……..62
Fig. 31 - The percentage (%) P3HT-fullerene number of contacts for ratio 1:0.8 of 15x15x80 box……..64
Fig. 32 - The percentage (%) P3HT-fullerene number of contacts for ratio 1:0.8 of 20x20x80 box……..65
Fig. 33 - The percentage (%) P3HT-fullerene number of contacts for ratio 1:0.8 of 20x20x80 box (second batch).....66
Fig. 34 - Spatial discretization scheme (a) original morphology (b) after remapping……..67

List of Abbreviations

AA	All-atom
BHJ	Bulk heterojunction
CB	Chlorobenzene
CBS	Complete basis set
CCSD	Coupled-cluster with single and double excitations
CCSD(T)	Coupled-cluster with single and double excitations, and third order perturbation theory corrections
CG	Coarse-grained
DFT	Density Functional Theory
DMB	1,3-dimethoxybenzene
EA	Electron Affinity
GGA	Generalized gradient approximation
GGE	Generalized gradient exchange
GROMACS	GROningen Machine for Chemical Simulations
HF	Hartree-Fock
h-GGA	Hybrid-generalized-gradient approximation
hh-GGA	Double-hybrid generalized-gradient approximation
hm-GGA	Hybrid-meta-generalized-gradient approximation
h-NGA	Hybrid-nonseparable-gradient-approximation
IP	Ionization Potential
KS	Kohn-Sham
LDA	Local density approximation
MaxE	Maximum error
MD	Molecular dynamics
m-GGA	Meta-generalized-gradient approximation
MP2	Second-Order Many-Body Perturbation
MPPT	Møller-Plesset Perturbation theory
MSE	Mean signed error
MUE	Mean unsigned error
NGA	Nonseparable gradient approximation
OPV	Organic photovoltaics
P3HT	poly(3-hexylthiophene)
PCBM	[6,6]-phenyl-C61-butyric acid methyl ester

PDB	Protein Data Bank
RDF	Radial distribution function
RMS	Root mean Square
RPA	Random-phase approximation
SD	Standard deviation
SE	Standard error
ZPE	Zero point energy

I. Benchmarking of Density Functionals for the Electron Affinity of the Fe(III)/Fe(II) Redox Pair

1. Introduction

Transition-metal complexes play an essential role in biological systems. Several enzymes require metal ions as cofactors for their biological functions. Based on the Protein Data Bank (PDB), there are almost 40% of proteins in which a metal is present. There are at least ten metals, including Na, Mg, Ca, Mn, Fe, Co, Ni, Cu and Zn, that may interact with one or more amino acid residues of protein molecules [1]. Many enzymatic reactions require metal ion cofactors for the redox catalysis or electron transfer reactions. The metal ion cofactor interacts with enzymes mostly through the coordination of amino acid side chains, for instance, the sulphur of cysteine (Cys), the imidazole nitrogen of histidine (His) or the carboxylate oxygen of aspartic and glutamic acids (Asp and Glu) [1, 2]. Metal cations usually interact with donor atoms in the functional groups of amino acid side chains.

Tamames and Ramos [3] have studied the occurrence of metal elements in proteins based on the PDB. Their study concluded that iron (Fe) shows the biggest number of occurrences in the different structures of proteins and enzymes. Iron is one of the most abundant transition metals on earth and essential for all living organisms. One of the important characteristics of iron (Fe) is its two oxidation states, Fe (II) and Fe (III). Iron may also exist in the redox IV state, which is transitory and can happen along catalysis. In enzymatic processes, the iron cofactor usually presents these two redox states (Fe (II) and Fe (III)), which allow for its participation in the catalysis of redox reactions [4]. The presence of multiple coordination and oxidation states has helped iron to become a popular cofactor to many biochemical reactions. It is involved in oxygen transport, electron transport and many other reactions [5]. Metalloproteins with an iron cluster are involved in electron transfer in redox reactions. Iron is essential regarding O_2 transport or electron transfer processes in cytochrome. In addition, iron-sulphur cluster and cupredoxins are necessary for the reduction potential reactions in biological systems [6]. Because of the important role played in biological processes, in

particular in redox reactions, it is essential the study of molecular redox properties of iron.

It is necessary to understand the key molecular properties accompanying redox reactions. The reduction potential is a direct assessment for understanding the thermodynamics of oxidation-reduction half reactions [7]. The study for redox properties will improve the understanding of important molecular features in the reduction potential. The reduction potential provides information about the equilibrium constant or free energy change for electrochemical half-reactions, which will play a role as the reduction part in a redox reaction.

In a real system, the complicated mechanisms involved in the biochemical systems cause the experimental measurements of reduction potentials very difficult to perform. Therefore, computational chemistry is a good solution for studying the properties of redox reactions. The calculation of potential energy surfaces allows for the calculation of free energies and redox potentials [7]. Most of the redox reactions occur in condensed phase, and there are several approaches for calculating the reduction potential. One method is based on thermodynamics cycles, which combines the gas-phase energetics with solvation free energies of the products and reactants. The general schematic representation of gas-phase and solution phase reactions can be seen in Figure 1.

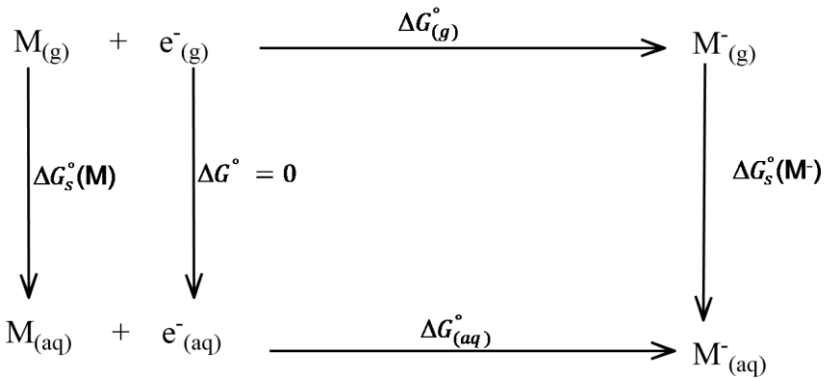


Fig 1. Thermodynamic cycle for the calculation of Gibbs free energies of one-electron reduction process [7, 8]

The explanation of the Gibbs free energy change, shown in Figure 1, involves a basic definition of molecular properties. The electron affinity (EA) is defined as the change in energy when a neutral atom gains an electron [9]. The Gibbs free energy of the gas-phase reaction, ΔG_g° , can be defined as the subtraction of $G^\circ(M^-)$ and $G^\circ(M)$ i.e. [7]:

$$\Delta G_g^\circ = -EA\,(M) + \Delta\Delta G_{therm} \quad (1)$$

The change in Zero Point Energy (ZPE) is included in the EA for eq (1). However, we can also separate it according to:

$$\Delta G_g^\circ = -\Delta E_{elec}\,(M) + ZPE + \Delta\Delta G_{therm} \qquad (2)$$

In which E_{elec} is the electronic energy for species M and M^-. In addition, ΔE_{elec} is the difference in the electronic energy between the two species of (M/M^-) and by far the dominant factor for calculating the Gibbs free energy of gas phase reaction. $\Delta\Delta G_{therm}$ is the thermal contribution to the free energy. In this work, we will focus in calculating the ΔE_{elec} on several iron complexes. We do not calculate $\Delta\Delta G_{therm}$ and the zero point ZPE because they are not the focus of our study. Additionally, Arumugam [8] remarked even though $\Delta\Delta G_{therm}$ and ZPE correction improve the calculation of reduction potential, they do not justify the very expensive computational calculations.

In recent years, theoretical studies regarding transition metals in biochemical systems are increasing very fast. This also reveals a series of challenging problems, which need to be tackled by theoretical chemists. It is important to find an accurate way for modelling transition-metal complexes that can mimic real systems in enzymes or proteins. The complexity of transition-metal complexes in biological systems requires high-accuracy methods for describing the systems [10, 11]. However, in contrast with the necessity of high-level quantum chemical levels for describing transition metal complexes, there are few studies that rely on the use of high-level accurate coupled cluster methods. Other consequences using high accuracy methods are the computational cost and the computational time. Therefore, high-level quantum chemical studies are not applicable to large systems, such as enzymes and proteins. There is an urgency for finding low-cost computational methods for studying these systems. The most promising solution is the use of density functional theory (DFT) methods. DFT methods have been in the centre of interest of computational chemistry because of their reasonable accuracy and applicability in large model systems. DFT has low scaling on the system size (N^3). Therefore, DFT is the most promising methodology for studying macromolecular systems, such as enzymes and proteins.[12]

The most important properties correlated with reduction potentials are ionization potentials (IP) and electron affinities (EA). Calculation of IP and EA of species in the gas phase is important for further studies in solution-phase reduction potentials. Several studies have been conducted to evaluate the performance of different DFT methods for molecular properties of transition metal complexes including ionization potential (IP) and electron affinity (EA). The study of IP and EA properties is an important starting step for further studies of redox potentials in transition metal

complexes. Redox potentials require the simultaneous transfer of electrons from one ion to another ion, which needs IP and EA to predict how much energy is needed for transfer one (or more) electron(s). Theoretical studies have been reported regarding the calculation of IP for atoms and molecules. Uudsemaa and Tamm [13] have calculated aqueous M^{3+}/M^{2+} redox potentials for fourth-period transition metals using the DFT methodology in combination with the COSMO continuum model. Their theoretical models include two solvation spheres (18 explicit water molecules). They compare their results with experimental redox potential results. The redox potentials of the ions in aqueous solution were calculated with the thermodynamic cycle pointed out above. They remark that the most important part for calculating the redox potential is the IP. The computed IP of the corresponding ion-water cluster was calculated as the difference of energies for the optimised geometries of the respective oxidation state cation complexes. However, in their studies, they only used the functional BP86 and did not do the benchmarking for finding the most suitable density functional for their theoretical models. Their studies were able to reproduce the experimental redox potential with an average unsigned error (MUE) of 0.29 V (6.68 kcal/mol).

Riley *et al.* [12] applied DFT, Hartree-Fock and second-order many-body perturbation (MP2) methods in order to assess several molecular properties of small molecules containing atoms commonly found in biomolecules, such as proteins, in particular, the C, H, N, O, S and P atoms. IP and EA are two of the molecular properties that they studied. They proved that DFT could predict the IP and EA properties with good results when compared with post-Hartree-Fock methods. They also concluded that the hybrid meta-GGA functional B1B95 with aug-cc-pVDZ basis set yields error of 4.53 and 4.54 kcal/mol for predicting the IP and EA values, in comparison with the respective experimental values. Li *et al.* [14] report their studies about benchmark adiabatic ionization potential (AIP) calculations of M/M^{+}-NH_3 (M = Na, Al, Ga, In, Cu and Ag) complexes, using DFT and CCSD(T) methods. The AIP values of neutral complexes were compared with the available experimental data. Benchmarking studies were done only for DFT methods since benchmarking of CCSD(T) methods is computationally expensive. They have tested several GGA and hybrid GGA and they got an average of 0.1 eV (2.31 kcal/mol) for all the complexes studied. They noticed that one big obstacle to benchmarking studies of the systems involving transition metals is the lack of experimental data. Su *et al.* [9] conducted a study to examine the performance of DFT functionals for calculating IP and electron affinity (EA), for several atoms and molecules for second and third-period elements and not for the transition metal complexes. The IP and EA values from several DFT

functions were compared with the available experimental data. The best functionals for IE are B2PLYP and M06-2X (MUE 2.51 and 2.74 kcal/mol) and the best functionals for EA are B2PLYP and PBE (MUE 2.08 and 2.35 kcal/mol).

Due to the lack of experimental IP and EA values, it is more difficult to evaluate the accuracy of DFT methods for transition-metal complexes. Moreover, there is a large number of DFT functionals with different accurate performances. Therefore, the choice of a proper DFT functional for calculating the IP of a transition metal complex is crucial. Until now, it is rare to find the benchmarking of DFT studies for calculating molecular properties of transition metal complexes. Here, we have conducted a benchmarking study of DFT methods to calculate the electron affinity of a transition metal, in particular iron.

Our main goals are to provide a benchmarking study on ΔE_{elec} of iron in different model complexes that mimic the different coordination shells, typically found in proteins, as well as finding which DFT functional gives a better performance for calculating the ΔE_{elec} of iron complexes. In order to perform a benchmarking, we need reference values that can be obtained from experimental data or using high-level methods. Here, we chose our reference values from accurate high-level methods. We used two post-Hartree-Fock methods, *i.e.* the second Moller-Plesset perturbation theory (MP2) and, single and double coupled cluster theory with perturbative triple correction (CCSD(T)). The extrapolation of energy to the complete basis set (CBS) limit was obtained by extrapolating the Hartree-Fock (HF) and correlation energies using two different schemes - Truhlar and Helgaker schemes [19, 21].

2. Literature Review

We will introduce the concepts of two key molecular properties associated with the calculation of reduction potentials. They are the ionization potential (IP) and the electron affinity (EA) properties. Furthermore, we will explain the relation between a thermodynamics cycle and the calculation of redox potentials. Lastly, we will describe some computational methods used to determine IP, EA and the redox potential.

2.1 Ionization potential (IP) and electron affinity (EA)

The ionization potential can be described as the energy that is needed for removing an electron from an atom or molecule. The adiabatic ionization potential (AIP) is defined as the energy required to form an atomic or molecular cation by releasing an

electron from the ground state of the neutral system in the gas phase. The first three ionization potentials are defined as the following:

First Ionization Potential $M \longrightarrow M^{+} + e^{-}$ (3)

Second Ionization Potential $M^{+} \longrightarrow M^{2+} + e^{-}$ (4)

Third Ionization Potential $M^{2+} \longrightarrow M^{3+} + e^{-}$ (5)

The electron affinity (EA) is defined in a similar way to the adiabatic ionization energy. However, the EA is the symmetric energy change because it corresponds to the variation energy that occurs when a neutral atom or molecule gains an electron.

$$M_{(g)} \longrightarrow M^{+}_{(g)} + e^{-},\ \Delta H_0 = IP \quad (6)$$

$$M_{(g)} + e^{-} \longrightarrow M^{+}_{(g)},\ \Delta H_0 = -EA \quad (7)$$

Adiabatic IP and EA corresponded to the energy differences between the lowest energy states of the neutral species and the corresponding ionic ones [9]. Zhou *et al*. [15] have defined the adiabatic ionization potential as the energy difference when the molecule is geometrically optimised. IP and EA are fundamental properties for assessing the electron donating and accepting capabilities of systems engaged with any redox processes. Moreover, these properties have been employed for understanding the electron and proton transfers that occur in the gas phase or in condensed phase [7, 9].

2.2 Computation of the Reduction Potential

In this section, we explain the related fundamental concepts for understanding the calculation of reduction potential. In electrochemical thermodynamics, there is a chemical potential (μ) for a pure substance and partial molar component of the solution [7]. For a species M in solution,

$$\mu_M = \mu_M^{\circ} + RT\ ln\left(\frac{\gamma C}{C^{\circ}}\right) = \mu_M^{\circ} + RT \ln(a) \quad (8)$$

where C is the concentration and the small subscript circle refers the value of the quantity in the standard state. a and γ are the activity and activity coefficient. In the gas phase, the usual standard state is an ideal gas at pressure of 1 bar, whereas for solutes in the liquid phase, the standard state is an ideal solution at the concentration of one molar (1 mol per liter solution) or 1 molal (1 mol of solute per kg solvent). We also introduce the dimensionless activity coefficients γ_i, which are defined as

$$a_i = \gamma_i \frac{C_i}{C_i^\circ} \quad (9)$$

If we apply Eq (7) to the reaction below:

$$\mathrm{Ox} + e^- \longrightarrow \mathrm{Red}$$

The free molar energy of reaction is written as

$$\Delta G = \Delta G^\circ + RT\, ln(Q) = \Delta G^\circ + RT \ln(\frac{a_{Red}}{a_{Ox}}) \quad (10)$$

where Q is the dimensionless reaction quotient. The free energy can be expressed in terms of electrode potential *E* of half-cell [8] as

$$\Delta G = -nFE \quad (11)$$

F is Faraday constant (96485 C/mol) and *n* is the number of electrons in the half-reaction. When we combine with Eq 8 it results in the Nernst equation [7]:

$$E = E^\circ + \frac{RT}{F} \ln\left(\frac{a_{Red}}{a_{Ox}}\right) \quad (12)$$

The Gibbs free energy difference for a half-reaction is the measurement required for calculating the standard reduction potential. Theoretical calculation of the reduction potential is performed for a half-cell reaction. The Gibbs energy difference of gas phase reaction in Figure 1 is the electron affinity of M and the thermal contribution to the free energy ($\Delta\Delta G_{therm}$). The Gibbs free energy of the gas-phase reaction, ΔG_g° can be defined as the subtraction of $G^\circ(M^-)$ and $G^\circ(M)$, and we will get the definition below:

$$\begin{aligned}
\Delta G_g^\circ &= G^\circ(M^-) - G^\circ(M) \\
&= [U_e(M^-) + ZPE\,(M^-) + \Delta G_{therm}(M^-)] - [U_e(M) + ZPE(M) + \Delta G_{therm}(M)] \\
&= -EA\,(M) + [\Delta G_{therm}\,(M^-) - \Delta G_{therm}\,(M)] \\
&= -EA\,(M) + \Delta\Delta G_{therm} \\
&= -\Delta E_{elec}\,(M^-/M) + ZPE + \Delta\Delta G_{therm}
\end{aligned} \quad (13)$$

U_e refers to the Born-Oppenheimer equilibrium potential energy. ZPE is the vibrational zero point energy and ΔG_{therm} is the thermal contribution to the free energy. The thermal contribution consists of free energy due to multiple conformations, rotation, vibrational and electronic excitation [7]. The free energy of the electron in Figure 1 is zero because in gaseous ion energetics, the free energy of a free electron is assigned as zero in the gas and solution phase [8]. In addition, we include the definition

of $-\Delta E_{elec}\ (M^{-}/M)$, which refers to the difference in the electronic energy between two species of M and by far the dominant factor for calculating the Gibbs energy change of gas phase reaction. In this work, we will focus on the calculation of the $-\Delta E_{elec}$ in different iron complexes.

From Figure 1, we can write the thermodynamics cycle for calculating the Gibbs free energy of one-electron reduction process for iron as below:

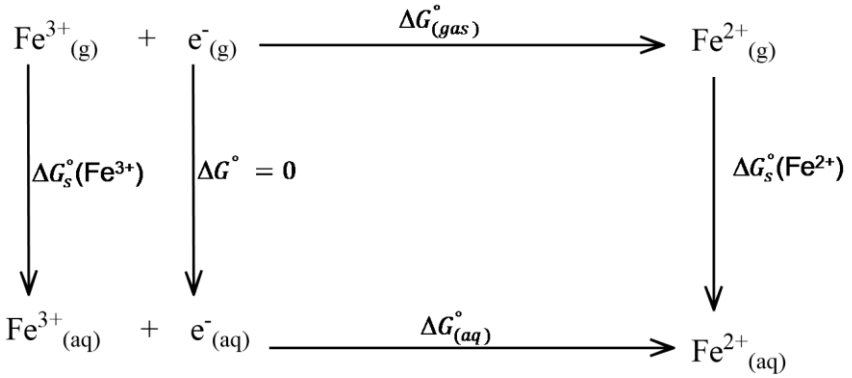


Fig. 2 Thermodynamics cycle for the calculation of the absolute and relative reduction potentials for the Fe (III)/Fe(II) redox pair.

In line with the definition of the Gibbs energy change in the gas phase, we defined $\Delta E'_{el}$ as the electronic energy from the subtraction of the electronic energies for the Fe(II) and Fe(III) complexes in our systems. $E_{el-Fe(II)}$ is described as the single point electronic energy calculation for different Fe (II) complexes and $E_{el-Fe(III)}$ electronic energy for different Fe (III) complexes.

$$\Delta E'_{el} = E_{el-Fe(II)} - E_{el-Fe(III)} \quad (14)$$

In our work, we will focus in the calculation of $\Delta E'_{el}$ for different small systems of iron complexes, in which the metal is coordinated with different ligands. The different metal complexes represent the most common amino acid side chains that bind to Fe(II) and Fe (III) in proteins and enzymes.

2.3 Methods

2.3.1 Moller-Plesset Perturbation Theory

The mathematical and computational techniques have been used to reach the HF limit and to solve the HF equation with the equivalent of an infinite basis set. When the HF limit is reached, the energy error associated with the HF approximation for a specific system is called electron correlation energy [16], and can be defined as

$$E_{corr} = E - E_{HF} \quad (15)$$

E is the true energy and E_{HF} is the system energy at the HF limit.

The HF theory makes the fundamental approximation that each electron moves in a static electric field created by all of the other electrons. There are two different electron correlations, which are the dynamic electron correlation and the static electron correlation. The dynamic contribution is referred as the instant correlation between electrons, for instance, the electrons that occupy the same spatial orbital. Whereas the static contribution is associated with electrons that avoid each other, such as those occupying different spatial orbitals [16, 17]. This contribution becomes important for systems where different configurations have similar energies.

In this section, we will discuss, in particular, the Moller-Plesset Perturbation Theory, which is one of several post-Hartree-Fock *ab initio* methods that include the electron correlation effects. We will start by explaining general principles of the perturbation theory [16]. The Hartree-Fock wavefunction Ψ_0 and energy E_0 are the approximate solution (eigenfunction and eigenvalue) to the exact Hamiltonian eigenvalue problem or the Schrödinger electronic wavefunction. However, the HF wavefunction and energy are the exact solution for the HF Hamiltonian H_0 eigenvalue problem. If we assume the HF wavefunction Ψ_0 and energy E_0 lie near the exact wave function Ψ and energy E, we can write the exact Hamiltonian operator [16] as

$$H = H_0 + \lambda V \quad (16)$$

Where V is a perturbation operator and λ is a dimensionless parameter.

If we expand the exact wavefunction and energy in terms of the HF wave function and energy we get

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \lambda^3 E^{(3)} + \cdots \quad (17)$$

$$\Psi = \Psi_0 + \Psi^{(1)} + \lambda^2 \Psi^{(2)} + \lambda^3 \Psi^{(3)} + \cdots \quad (18)$$

The substitution of these expansions into the Schrödinger equation according to the power of λ yields:

$$H_0 \Psi_0 = E^{(0)} \Psi_0 \quad (19)$$

$$H_0 \Psi^{(1)} + V \Psi_0 = E^{(0)} \Psi^{(1)} + E^{(1)} \Psi_0 \quad (20)$$

$$H_0 \Psi^{(2)} + V \Psi^{(1)} = E^{(0)} \Psi^{(2)} + E^{(2)} \Psi_0 \quad (21)$$

Multiplying each of the equations by Ψ_0 and integrating over all space, yields the expression for the n-th-order (MPn) energy:

$$E^{(0)} = \langle \Psi_0 | H_0 | \Psi_0 \rangle \quad (22\text{ a.})$$

$$E^{(1)} = \langle \Psi_0 | V | \Psi_0 \rangle \quad (22\text{ b.})$$

$$E^{(2)} = \langle \Psi_0 | V | \Psi^{(1)} \rangle \quad (22\text{ c.})$$

Then, the Hartree energy is the sum of the zeroth and the first order energies

$$E_0 = \langle \Psi_0 | H_0 + V | \Psi_0 \rangle \quad (23\text{ a.})$$

$$E_0 = E^{(0)} + E^{(1)} \quad (23\text{ b.})$$

The correlation energy is written as

$$E_{corr} = E_0^{(2)} + E_0^{(3)} + E_0^{(4)} + \cdots \quad (24)$$

The first term is the the MP2 energy.

The calculation of the MP2 energy is fairly inexpensive and can be efficiently evaluated. The scaling behaviour of the MP2 method is N^5 where N is the number of basis functions [17]. MP2 typically accounts for 80-90% of the correlation energy and it is the most economical molecular orbital method for including electron correlation [16].

2.3.2 Single and Double Coupled Cluster Theory With Perturbative Triple Correction (CCSD(T))

Another higher level theory for estimating electron correlation energy is the coupled-cluster (CC) theory. In the CC theory, the full configuration interaction (CI) wave function can be described [17] as:

$$\Psi = e^{\boldsymbol{T}} \Psi_{HF} \quad (25)$$

The cluster operator **T** is defined as

$$\mathrm{T} = \mathrm{T}_1 + \mathrm{T}_2 + \mathrm{T}_3 + \cdots + \mathrm{T}_n \quad (26)$$

Where n is the total number of electrons and the various T_i operators generate all possible determinants having i excitations from the reference state.

The scaling behaviour of CCSD is in the order of N^6 [16]. The inclusion of connected triple excitation for those arising with their own amplitudes from T_3 not the disconnected triples as product of T_1 and T_2, defines CCSDT. However, scaling as N^8 is computationally very costly, making it unfeasible for even the smallest molecules.

Several approaches have been proposed for estimating the effects of the connected triple excitations using perturbation theory. One of the most commonly used, within the coupled-cluster theory, is the CCSD(T) method. This method suggests a significantly improved augmented CCSD technique that can be obtained by using a triple formula, which results from interactions with both singles and doubles and analogous to what is used in the QCISD(T) method [18]. The (T) approach in general slightly overestimates the triples correction and does so by an amount about equal to the ignored quadruples, with a favourable cancellation of errors. Therefore, the CCSD(T) method is extremely effective and accurate for single reference calculation [17].

Accurate electronic energies of reactions can be achieved using calculations at high-level theory, for instance, CCSD(T) with a large basis set. However, due to the fact that a CCSD(T) calculation scales as the seventh power of the number of atoms in the system and largely increases in computational cost, the calculations using high-level theory should be restricted to relatively small molecular systems [7, 16, 17]. Another alternative to wavefunction theory (WFT) is the density functional theory (DFT), in which the computational work scales as N_{basis}^{3} or N_{basis}^{4} rather than N_{basis}^{7}, where N_{basis} is the number of atoms in the system.

2.3.3 Extrapolation to the Complete-Basis-Set

The basis set truncation and approximation in the correlation treatment make it almost impossible to compute the full correlation energy. There are several considerations for using electronic structure calculations with extrapolation-limit basis sets. For a given choice of basis set and level of theory, we can employ extrapolations to complete-basis-set (CBS) energies. The solution of the HF equations with a complete basis-set is called as HF limit. For this purpose, we can use basis set cc-pVXZ (X = D, T, Q). Increasing in size will be in a consistent pattern with each increment of X [17, 19]. The extrapolation to complete basis set is important because of the slow convergence of correlated calculations to the limit of a complete basis set.

Several attempts for extrapolating have been done for instance Halkier *et al.* [20] obtained results within a few tenths of kcal/mol of the basis-set limit for a CCSD(T) calculation by extrapolating from cc-pCV5Z and cc-pCV6Z calculations. However, their results are computationally costly. Truhlar [19] has developed a more economical basis set extrapolation of correlated electronic structure based on correlation-consistent polarised double and triple zeta basis sets. Extrapolations from cc-pVDZ and cc-pVTZ are useful for the application to large molecular systems. In his method, rather than

minimise the mean unsigned error (MUE), he minimises the root-mean-square (RMS) error of the extrapolated exponent. The extrapolations have been done separately for Hartree-Fock and correlation energies. The total energy is the sum of HF and the correlation parts:

$$E^{tot} = E^{HF} + E^{cor} \tag{27}$$

The components of the energy are assumed to approach their basis-set limits by power laws:

$$E_X^{HF} = E_\infty^{HF} + A^{HF} X^{-\alpha} \tag{28}$$

$$E_X^{HF} = E_\infty^{HF} + A^{HF} X^{-\beta}$$

$$E_\infty^{tot} = \frac{3^\alpha}{3^\alpha - 2^\alpha} E_3^{HF} - \frac{2^\alpha}{3^\alpha - 2^\alpha} E_2^{HF} + \frac{3^\beta}{3^\beta - 2^\beta} E_3^{cor} - \frac{2^\beta}{3^\beta - 2^\beta} E_2^{cor} \tag{29}$$

The optimisation of exponents yielded values of alpha α = 3.4, β_{MP2}= 2.2 and $\beta_{MP2} = \beta_{CCSD(T)}$ = 2.4. Truhlar's method [19] works economically well and shows that the extrapolated calculation is only 5% more expensive than performing only the cc-pVTZ calculation. In addition, the RMS error of the method is small. This method is an *ab initio* method because the data used for parameterization are all *ab initio* data.

Helgaker *et al.* [21] presented a systematic calculation for investigating the basis-set convergence for HF, MP2 and CCSD(T) using the correlation-consistent basis sets of Dunning. They remarked that we need to distinguish between convergence of the HF energy and correlation energy. The HF method allows for an exact treatment within a finite one-electron basis set, whereas the correlation energy cannot be treated exactly in any finite-dimensional basis set. Therefore, the convergence character between these contributions is different. For extrapolating the HF and correlation energies, Helgaker *et al* [21] used analytical forms, which explain the behaviour and allow for accurate fits of the calculated energies with few parameters. The energy lowering along the cc-pVXZ decreases approximately geometrically in correlated calculations. In their studies, they used a water molecule and plotted the cc-pVXZ energies with the exponential fit using the equation below:

$$E_{SCF} = a + b \exp(-cX) \tag{30}$$

This exponential fit is found to be excellent for all the basis sets and give an indication that SCF decreases in a geometrical way for correlation-consistent basis sets.

The extrapolation of correlation energy defined by Helgaker *et al* [21] follows the equation below

$$E_{corr} = a + b\,X^{-3} \quad (31)$$

The calculation of energies at the cc-p(C)VXZ level, using triple-zeta and higher level basis sets, are fitted to eq (31).

2.3.4 Density Functional Theory (DFT)

The fundamental theorem of DFT is the Hohenberg-Kohn theorem, which states that the external potential is a functional of the ground-state density. The interaction between electrons is described using density in an observable 3D space [8]. An important step for applying DFT in real systems was taken in 1965 when Kohn-Sham published a Kohn-Sham (KS) equation, which is derived from the Hohenberg Kohn theorem. In DFT, the general form of energy as a function of density is described [22, 23] as

$$E[\rho(r)] = T_e[\rho(r)] + V_{ne}[\rho(r)] + V_{ee}[\rho(r)] + E_{xc}[\rho(r)] \quad (32)$$

$T_e[\rho(r)]$ is the kinetic energy of electron, $V_{ne}[\rho(r)]$ is the nuclear-electron interaction, $V_{ee}[\rho(r)]$ is the Coulomb repulsion, and lastly $E_{xc}[\rho(r)]$ is the exchange-correlation functional. The DFT wave function is constructed in a different way than in HF, and the resulting orbitals are referred as Kohn-Sham (KS) orbitals. The main difference between the self-consistent field (SCF) and KS-DFT approaches is the exchange-correlation functional. The dynamics electron correlation can be included through exchange-correlation in Density Functionals (DF). This step is one of the most important aspects of developing DF, which accurately calculate the electronic structure and properties [24]. The priority for developing better DF will be to find more specifically the exchange-correlation functional $E_{xc}[\rho(r)]$, which works for the universal chemistry and biological systems. We can classify the DFs into five major groups that consist of local density approximation (LDA), generalized gradient approximation (GGA), meta-GGA, hybrid functionals, and random phase approximation [25, 26].

The hierarchy of density functional approximation can be described with Jacob's ladder [26]. The earth is Hartree approximation ($E_{xc} = 0$) and heaven is the high chemical accuracy. Each rung is built from the elements of the lower rungs. Each of the rungs suffices certain exact constraints and when a form for certain rung want to advance to the next rung then it should have functionals based on the previous lower rung. The accuracy of the higher rungs is complemented by the lower rung [25]. The lowest rung of ladder is the LDA, which employs only the local density functionals. The second rung of the ladder is the GGA functionals that add a second component, the

gradient of density. The third rung corresponds to the meta-GGA functionals, which add the orbital kinetic energy density as the third component[27]. The first three rungs are semi local functionals, which are computationally efficient. The fourth and fifth rungs are non-local functionals that are computationally more expensive because of requirements at least double integration over three-dimensional space. However, they are more accurate [26]. The fourth rung corresponds to hybrid functionals, which implement an exact exchange component from HF theory. On the fifth rung, there is random-phase approximation such as functionals that use the unoccupied Kohn-Sham orbitals. In that rung, DFT starts to have similarity to many-body theory.

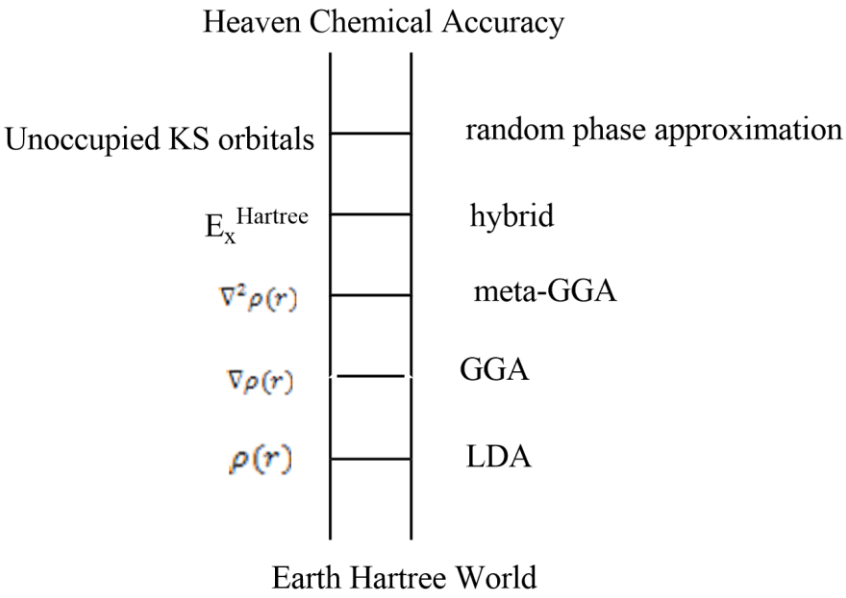


Fig. 3 - Jacob's Ladder of density functional approximation [26]

The first approximation introduced was the local density approximation (LDA). LDA simplifies the electron-electron interactions by including the interaction between electrons and the charge density of the other electrons. [28]. The generalisation of LDA that allows different spatial orbital for electrons with the opposite spin is known as local spin-density approximation [24]. The improvements in LDA and LSDA lead to the discovery of generalized gradient approximation (GGA). GGA functionals include the gradient, which leading the functionals to contain their function of spin densities plus their gradient [25, 26]. GGA functionals are often called semilocal functionals because some of the nonlocal effects are integrated for determining a contribution of the total exchange-correlation energy. Total exchange-correlation energy in GGA is split into exchange and correlation parts.

Meta-GGA functionals include the electron density $\rho(r)$, its gradient $\nabla\rho(r)$ and Laplacian of density $\nabla^2\rho(r)$ and the orbital kinetic energy (semi-local interactions),

which make a better improvement in functionals. Becke [29] proposed the mix of GGA with the exact exchange, which leading to the discovery of hybrid functionals. These are functionals that include an exact exchange component from HF theory. Perdew [26] explains the hybrid functionals using the PBE functional. The PBE0 is a hybrid functional that combines the full PBE correlation and 25% of a semi empirical parameter of exact exchange with a complementary fraction of PBE-GGA exchange. This functional is more accurate than PBE for calculating ion of atomization energy and surface energies. The fifth-rung functionals are the random-phase approximation (RPA-like functionals). The RPA functionals use Kohn-Sham and not Hartree orbitals. RPA can treat long-range correlation exactly, and therefore, compatible with the exact exchange. RPA functionals are poorly performed for short-range correlation. However, short-range correlation can be included into RPA-like equations through a local field factor [25].

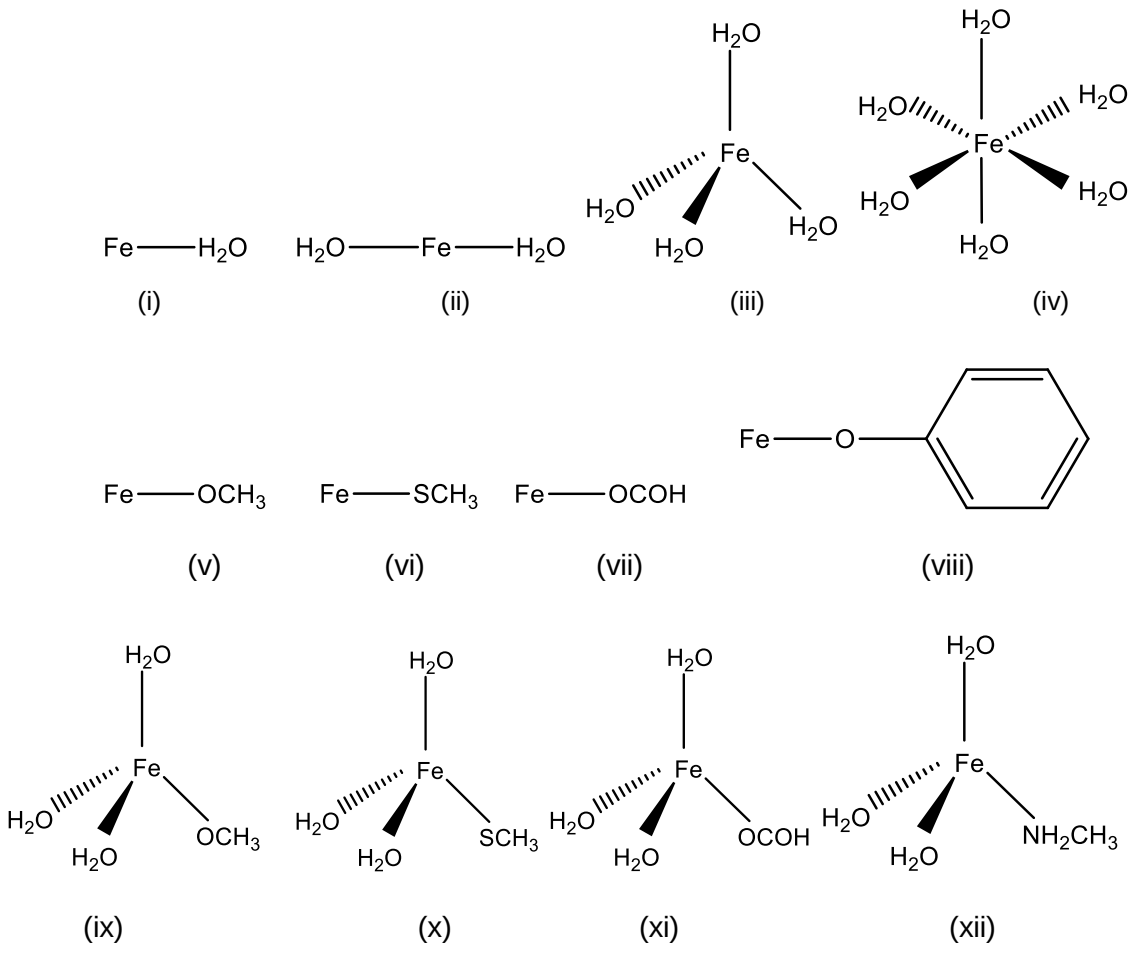
In addition to the explanation above, there are several recent developments in density functionals. One of these ideas is the range-separated functionals. Range-separated functionals divide the separation of electrons into two parts, one long-range and the other short-range and treating the two parts with different functionals [28]. The way to construct range separated functionals is mixing the long-range HF with short-range functionals, and forming simple long-range corrected DFs [24]. Another development is through a fitting process. The challenges are maintained because it is very hard to construct an accurate form of exchange-correlation for molecular systems. One practical solution is taking the available experimental information to help in the determination and tests of the functional [22, 28]. We can apply some forms of parameterization of the functional to the chosen sets of experimental data [28]. Truhlar's group has combined and extended the ideas with parameterization to a large number of chemically important species and give results for useful functionals such as M06-L, M06, M06-2X and M06-HF [30].

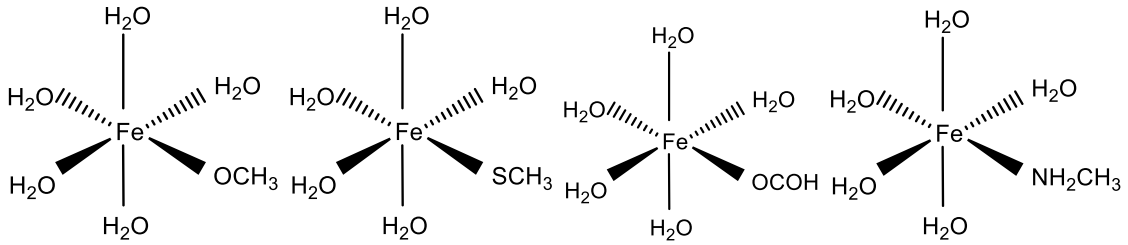
Even though there are several of density functionals present nowadays, it is still hard to find the functional that is suitable for all chemical and biological systems. In addition, choosing the right functional for a particular system also become more difficult. Therefore, in our studies, we attempt to perform the benchmarking studies of several density functionals and comparing their results with calculations from high-level theory methods, for instance, CCSD(T) with extrapolation to complete basis set.

3. Computational Details

3.1 Model Systems

We have used 16 model systems in our study. These models can have different types of geometries, such as linear, square planar or tetrahedral and octahedral. The two coordination numbers have a linear shape; the four coordination number adopts the tetrahedral or square planar geometry and the six coordination number shows an octahedral geometry. Some of the model systems are composed by different ligands, for instance, CH_3O^-, $HCOO^-$, CH_3S^- and $C_6H_5O^-$. These ligands represent the side chains of some amino acids. For instance, CH_3O^- represents the serine (Ser) side chain, $HCOO^-$ represents the glutamate (Glu) and aspartate (Asp) side chain, CH_3S^- represents the cysteine (Cys) side chain and $C_6H_5O^-$ represents the tyrosine (Tyr) side chain. All of the calculations have been performed with the Gaussian09 package program [31]. Our models are illustrated in figure 4.





(xiii) (xiv) (xv) (xvi)

Fig 4. Representation of the different complexes with Fe(II) and Fe(III) studied:
(i) $Fe(H_2O)$ (ii) $Fe(H_2O)_2$ (iii) $Fe(H_2O)_4$ (iv) $Fe(H_2O)_6$ (v) $Fe(OCH_3)$ (vi) $Fe(SCH_3)$ (vii) $Fe(COOH)$ (viii) $Fe(C_6H_5O)$ (ix) $Fe(H_2O)_3(OCH_3)$ (x) $Fe(H_2O)_3(SCH_3)$ (xi) $Fe(H_2O)_3(COOH)$ (xii) $Fe(H_2O)_3(NH_2CH_3)$ (xiii) $Fe(H_2O)_5(OCH_3)$ (xiv) $Fe(H_2O)_5(SCH_3)$ (xv) $Fe(H_2O)_5(COOH)$ (xvi) $Fe(H_2O)_5(NH_2CH_3)$

3.2 Optimisation of Structures

The geometries of all complexes were optimised at the MP2/6-311+G(d,p) level of theory. The geometry optimisations were carried out at the MP2 theoretical level to avoid bias toward any functional and also because it has proven to give good results as far as the optimisation of geometries is concerned [32]. The geometry optimisation calculations were carried out separately for the Fe(II) and Fe(III) complexes. There are two different spin states for both Fe(II) and Fe(III). For Fe(II), the high spin multiplicity is five and the low spin multiplicity is one. The high spin multiplicity of Fe(III) is six, while its low spin multiplicity is two. The spin states which gave the lowest energy were used for further computation using a higher level theory [11] (please see Supporting Information for calculated multiplicities for each complex). It turned out that high spin multiplicity (five for Fe(II) and six for Fe(III)) gave significantly lower energies (-73.38 kcal/mol and -67.38 kcal/mol) when comparing with low multiplicities, as expected. All the calculations were performed using the Gaussian09 program.

3.3 Calculation of the electronic energy E'_{el}

The reference data were obtained from the calculation at a higher level of theory at MP2 and at CCSD(T) higher level of theory. For each optimised structure, we have carried out single point energy calculations in vacuum at MP2 level using Dunning basis set aug-cc-pVXZ (X = 2, 3, 4). Because of the high computational cost, we only performed calculations for CCSD(T) using the aug-cc-pVDZ level of theory. In order to determine the energy at CCSD(T)/CBS level of theory, we employed two extrapolation methods, which were developed by Truhlar (scheme 1) and Helgaker (scheme 2). We used single point (SP) energy calculations from MP2/aug-cc-pVXZ (X = 2, 3, 4) level of

theory to extrapolate for the MP2/CBS level. The total energy at the CBS limit ($E_{MP2/CBS}$) was defined as the sum of the CBS Hartree-Fock energy ($E_{HF/CBS}$) and the CBS correlation energy ($E_{corr/CBS}$)(eq.33).

$$E_{MP2/CBS} = E_{HF/CBS} + E_{cor/CBS} \quad (33)$$

The extrapolation schemes of HF energy and correlation energy were determined separately because the HF energy converges faster than the correlation energy. To obtain Truhlar [19] extrapolation energy, we employed a single point energy using the aug-cc-pVDZ and aug-cc-pVTZ Dunning basis sets (eq 29).

$$E_{MP2/CBS}^{Truhlar} = \frac{3^{\alpha}}{3^{\alpha} - 2^{\alpha}} E_{HF/TZ} - \frac{2^{\alpha}}{3^{\alpha} - 2^{\alpha}} E_{HF/DZ} + \frac{3^{\beta}}{3^{\beta} - 2^{\beta}} E_{corr/TZ} - \frac{2^{\beta}}{3^{\beta} - 2^{\beta}} E_{corr/DZ}$$

The parameters α and β are 4.93 and 2.13, respectively. $E_{corr/DZ}$ and $E_{corr/TZ}$ were calculated separately. $E_{corr/DZ}$ was obtained from subtraction between $E_{MP2/DZ}$ and $E_{HF/DZ}$. Similarly, $E_{corr/TZ}$ was obtained from subtraction of $E_{MP2/TZ}$ and $E_{HF/TZ}$.

We also calculated the extrapolation energy to the CBS limit using Helgaker [21] extrapolation scheme. In the Helgaker's scheme, we performed the exponential fit of HF using the equation below.

$$y = c + a \exp(bX) \quad (34)$$

y value corresponds to $E_{HF/XZ}$ and X value corresponds to 2, 3, 4. After the exponential fitting process, we will get the c value, which corresponds to the $E_{HF/CBS}^{Helgaker}$ value. We did the extrapolation of the correlation energy using the equation below

$$E_{corr/CBS} = \frac{4^3 E_{corr/QZ} - 3^3 E_{corr/TZ}}{4^3 - 3^3} \quad (35)$$

$E_{corr/QZ}$ was obtained from subtraction of $E_{MP2/QZ}$ and $E_{HF/QZ}$. The extrapolation correlation energy was calculated using aug-cc-pVTZ and aug-cc-pVQZ, instead of using aug-cc-pVDZ and aug-cc-pVTZ because they have greater accuracy in extrapolated results [21]. The final CBS energy is defined as the sum of $E_{HF/CBS}$, which we get from exponential fit in eq (32), and the CBS correlation energy.

$$E_{MP2/CBS}^{Helgaker} = E_{HF/CBS} + E_{corr/CBS}^{Helgaker} \quad (36)$$

The difference between the CCSD(T) and CBS energies ($\Delta E_{(CCSD(T)-MP2)}$) is assumed to have small basis set dependence [33]. The final CCSD(T)/CBS energy will consist of the difference between the CCSD(T) and CBS energies ($\Delta E_{(CCSD(T)-MP2)}$) using the

aug-cc-pVDZ basis set and the MP2/CBS energy from the Helgaker extrapolation scheme.

$$E_{CCSD(T)/CBS} = E_{MP2/CBS}^{Helgaker} + \Delta E_{(CCSD(T)-MP2)/DZ} \quad (37)$$

The reference value of $\Delta E_{el}'$ for the DFT benchmarking is defined as eq (38)

$$\Delta E_{el}' = E_{el-Fe(II)\,CCSD(T)/CBS} - E_{el-Fe(III)CCSDT(CBS)} \quad (38)$$

3.4 Benchmarking of DFT functional

E_{el}' has been calculated using DFT for determining which density functionals have better performance for our systems. We have used 45 density functionals including LDA, GGA, meta-GGA, hybrid GGA, hybrid meta-GGA and range-separated density functionals. All of the benchmarking calculations were carried out using the 6-311++g(2df,2p) basis set. E_{el}' can be obtained as the difference between the single point energy calculations of optimized Fe(II) and Fe(III) complex structures.

Table 1. List of DFT Functionals and basis sets used in this work.

Functional			**Basis Set**
LDA	hm-GGA	Double hybrid functional	Pople type
SVWN [34]	B1B95 [35]	B2GPPLYP [36]	6-311+G(d,p) (for geometry optimization)
SPW91[37]	BB1K [38]	B2PLYP [39]	6-311++g(2df,2p) (for DFT calculation)
	BMK [40]	DSD-BLYP [41]	Correlation consistent
GGA	M05 [42]	MPW2PLYP [43]	aug-cc-pVDZ
BP86 [44-46]	M05 - 2X [47]		aug-cc-pVTZ
BPBE [44, 48]	M06 - 2X [30]		aug-cc-pVQZ
BPW91 [44, 49]	M06 [30]	NGA	
G96LYP [50, 51]	MPW1B95 [49, 52-54]	N12 [55]	
HCTH407 [56]	MPWB1K [53]		
OLYP [50, 57]	TPSSh [58]	Range separated HM-GGA	
OPL [59, 60]		M11 [61]	
	h-GGA		
m-GGA	B1LYP[50, 52]	Range-separated hm-NGA	
OTPSS [60]	B3LYP[50, 62]	MN12-SX [63]	
M06 – L [30]	B3P86[45, 62]	h-NGA	

VSXC [64]	B3PW91 [65]	MN12 – L [66]
TPSSTPSS [58]	BhandH [65]	
M11 – L [67]	MPW1K [68]	GGE
MPWB95 [49, 52, 54]	MPW1N [49, 54, 69]	OVWN5 [34, 70]
	PBE1PBE [71]	
Range separated functional	B97-1 [72]	
wB97XD [73]	B97-2 [74]	

4. Results and Discussion

4.1 Calculation of $\Delta E'_{el}$ at the MP2/CBS level

There are several methods which can be chosen for calculating chemical properties. Extrapolation to complete basis set (CBS) using high-level post-HF methods is one of the methods, which provides more accurate results. We have used high-level post-HF *ab initio* methods MP2 and CCSD(T), which were corrected for the basis set truncation error using basis set extrapolation to obtain reference values of CCSD(T)/CBS energies. Moreover, we compare the reference values with results of 45 density functionals to investigate the performance of various DFT functionals in predicting the electronic energy contribution, E'_{el}, of Fe(II) and Fe(III) for different systems. Table 2 shows a summary of $\Delta E'_{el}$ (Fe(II) - Fe(III)) quantities determined for different systems at the MP2 level for aug-pVDZ, aug-cc-PVTZ and aug-cc-pVQZ basis set.

Table 2. $\Delta E'_{el}$ (Fe(II) - Fe(III)) at the MP2 level using aug-cc-pVDZ, aug-cc-pVTZ and aug-cc-pVQZ basis sets

Model		$\Delta E'_{el}$ (Fe(II) - Fe(III)) (kcal/mol)		
		MP2/aug-cc-pVDZ	MP2/aug-cc-pVTZ	MP2/aug-cc-pVQZ
1	$Fe(H_2O)$	-596.6546	-599.1267	-601.7584
2	$Fe(H_2O)_2$	-532.2911	-534.6569	-537.1669
3	$Fe(H_2O)_4$	-436.5451	-438.9066	-441.2416
4	$Fe(H_2O)_6$	-381.3429	-383.5432	-385.8932
5	$Fe(OCH_3)$	-345.4137	-351.1872	-353.4012
6	Fe(OOCH)	-336.8480	-340.8949	-342.6472

7	$Fe(SCH_3)$	-316.8745	-320.0197	-321.5770
8	$Fe(C_6H_5O)$	-292.7158	-295.5639	-296.5444
9	$Fe((H_2O)_3(NH_2CH_3)$	-418.7232	-421.6494	-423.9953
10	$Fe(H_2O)_3(COOH)$	-300.7856	-303.4671	-305.5607
11	$Fe(H_2O)_3(OCH_3)$	-271.8355	-274.9105	-276.9332
12	$Fe(H_2O)_3(SCH_3)$	-258.7915	-261.2471	-262.7310
13	$Fe(H_2O)_5(NH_2CH_3)$	-370.4212	-373.1151	-375.4923
14	$Fe(H_2O)_5(COOH)$	-269.0788	-271.3712	-273.4479
15	$Fe(H_2O)_5(OCH_3)$	-246.0797	-248.7022	-250.6648
16	$Fe(H_2O)_5(SCH_3)$	-236.5793	-238.9616	-240.4693

Table 3 summarises the $\Delta E'_{el}$ (Fe(II) - Fe(III)) values at the MP2/CBS level, which were obtained using the Truhlar extrapolation scheme. Subsequently, we disclose the differences between $\Delta E'_{el}$ at the MP2/CBS and MP2/aug-cc-pVXZ (X= D, T, Q) levels. The results reveal that the extrapolated values $\Delta E'_{el}$ MP2/CBS are closer to quadruple-zeta than to triple-zeta basis sets. The values of mean signed error (MSE) and mean unsigned error (MUE) to the CBS value (1.47 kcal/mol for both) for quadruple-zeta basis set are smaller than for triple-zeta. The convergence of $\Delta E'_{el}$ with the basis set is slow because of the aug-cc-pVQZ results are still 1.47 kcal/mol from the CBS limit. Considering the MSE and MUE values between the three methods, we can conclude that the extrapolated values for these systems are more accurate than the quadrupole-zeta values.

Table 3. $\Delta E'_{el}$ (Fe(II) - Fe(III)) at MP2/CBS using Truhlar Extrapolation Scheme.

	Model	$\Delta E'_{el}$ (Fe(II) - Fe(III)) (kcal/mol)			
		MP2/CBS Truhlar	MP2/CBS - MP2/aug-cc-pVDZ	MP2/CBS - MP2/aug-cc-pVTZ	Truhlar MP2/CBS - MP2/aug-cc-pVQZ
1	$Fe(H_2O)$	-602.8805	-6.23	-3.75	-1.12
2	$Fe(H_2O)_2$	-538.3929	-6.10	-3.74	-1.23
3	$Fe(H_2O)_4$	-442.5902	-6.05	-3.68	-1.35
4	$Fe(H_2O)_6$	-386.9861	-5.64	-3.44	-1.09
5	$Fe(OCH_3)$	-355.4209	-10.01	-4.23	-2.02
6	Fe(OOCH)	-344.1277	-7.28	-3.23	-1.48
7	$Fe(SCH_3)$	-322.7887	-5.91	-2.77	-1.21
8	$Fe(C_6H_5O)$	-298.4559	-5.74	-2.89	-1.91
9	$Fe((H_2O)_3(NH_2CH_3)$	-425.3816	-6.66	-3.73	-1.39
10	$Fe(H_2O)_3(COOH)$	-307.3227	-6.54	-3.86	-1.76
11	$Fe(H_2O)_3(OCH_3)$	-279.0221	-7.19	-4.11	-2.09

12	$Fe(H_2O)_3(SCH_3)$	-263.8079	-5.02	-2.56	-1.08
13	$Fe(H_2O)_5(NH_2CH_3)$	-376.6431	-6.22	-3.53	-1.15
14	$Fe(H_2O)_5(COOH)$	-274.9743	-5.90	-3.60	-1.53
15	$Fe(H_2O)_5(OCH_3)$	-252.6365	-6.56	-3.93	-1.97
16	$Fe(H_2O)_5(SCH_3)$	-241.5428	-4.96	-2.58	-1.07
	MSE		-6.38	-3.48	-1.47
	MUE		6.38	3.48	1.47

Furthermore, the extrapolation values of $\Delta E'_{el}$ (Fe(II) - Fe(III)) at the MP2/CBS level using the Helgaker extrapolation scheme have been also calculated. This extrapolation involves the electronic energies at the MP2/aug-cc-pVTZ and aug-cc-pVQZ levels, as well as the HF energies that have been exponentially fitted using the equation (32). It is expected that the values obtained using the Helgaker scheme are more accurate since they use larger basis sets than in Truhlar scheme. Table 4 shows the MP2/CBS $\Delta E'_{el}$ obtained using the Helgaker method as well as the difference between the two methodologies.

Table 4. $\Delta E'_{el}$ (Fe(II) - Fe(III)) at the MP2/CBS level using the Helgaker Extrapolation Scheme and the difference between Truhlar and Helgaker Extrapolation Methods.

		$\Delta E'_{el}$ (Fe(II) - Fe(III)) (kcal/mol)			
	Model	Truhlar MP2/CBS	Helgaker MP2/CBS	Truhlar MP2/CBS - Helgaker MP2/CBS	Helgaker MP2/CBS - MP2/aug-cc-pVQZ
1	$Fe(H_2O)$	-602.8805	-602.2960	-0.58	-0.54
2	$Fe(H_2O)_2$	-538.3929	-539.5072	1.11	-2.34
3	$Fe(H_2O)_4$	-442.5902	-443.4756	0.89	-2.23
4	$Fe(H_2O)_6$	-386.9861	-388.1550	1.17	-2.26
5	$Fe(OCH_3)$	-355.4209	-354.2883	-1.13	-0.89
6	Fe(OOCH)	-344.1277	-343.9694	-0.16	-1.32
7	$Fe(SCH_3)$	-322.7887	-322.8267	0.04	-1.25
8	$Fe(C_6H_5O)$	-298.4559	-297.3581	-1.10	-0.81
9	$Fe((H_2O)_3(NH_2CH_3)$	-434.4150	-426.1135	0.73	-2.12
10	$Fe(H_2O)_3(COOH)$	-307.3227	-307.3763	0.05	-1.82
11	$Fe(H_2O)_3(OCH_3)$	-279.0221	-281.2145	2.19	-4.28
12	$Fe(H_2O)_3(SCH_3)$	-263.8079	-264.0199	0.21	-1.29
13	$Fe(H_2O)_5(NH_2CH_3)$	-376.6431	-426.1135	1.03	-2.18
14	$Fe(H_2O)_5(COOH)$	-274.9743	-275.4854	0.51	-2.04
15	$Fe(H_2O)_5(OCH_3)$	-252.6365	-252.3724	-0.26	-1.71
16	$Fe(H_2O)_5(SCH_3)$	-241.5428	-241.7543	0.21	-1.29

MSE	0.31	-1.77
MUE	0.71	1.77

The results show that the difference between Helgaker MP2/CBS and MP2/ aug-cc-pVQZ is significant, having MUE value of 1.77 kcal/mol. It can be concluded that it is not enough to use MP2/aug-cc-pVQZ and it is necessary to perform extrapolation to CBS. In addition, the difference between the $\Delta E'_{el}$ values at MP2/CBS level extrapolated using both methods is small, but meaningful, in particular considering the accuracy that is expected from the CBS extrapolation. The difference values obtained for the different model systems lie in the range between 0.20 – 1.17 kcal/mol. The only exception is the case of $Fe(H_2O)_3(OCH_3)$ that has the difference of 2.19 kcal/mol. The application of larger basis set in the Helgaker scheme improves the obtained MP2/CBS energies for all these model systems. Therefore, we decided to use the Helgaker extrapolation scheme value as our reference values.

4.2 Calculation of $\Delta E'_{el}$ at the CCSD(T)/CBS level

We obtained electronic energy contribution $\Delta E'_{el}$ (Fe(II) - Fe(III)) for CCSDT/CBS by adding the MP2/CBS energies with the energy difference between CCSD(T) and MP2 ($\Delta E_{CCSD(T)-MP2}$). Because of the high computational cost, here we only performed CCSD(T) calculations using aug-cc-pVDZ basis set. Jurečka and Hobza [33] have performed studies related to the energy difference between the CCSD(T) and MP2 ($\Delta E_{CCSD(T)-MP2}$) term and they concluded that it converges faster to the CBS than the MP2 and CCSD(T) energies. They remarked that aug-cc-pVDZ basis set yields values close to cc-pVTZ basis set with the largest difference less than 0.1 kcal/mol. In addition, Mládek *et al* [75] noted that the correlation energy term ($\Delta E_{CCSD(T)-MP2}$) is rather invariant to the number of basis functions from 6-31+g(d) to aug-cc-pVDZ with the average difference below 0.05 kcal/mol.

Table 5. $\Delta E'_{el}$ (Fe(II) - Fe(III)) at CCSD(T)/CBS, MP2/CBS and the difference between each method.

	Model	$\Delta E'_{el}$ (Fe(II) - Fe(III)) (kcal/mol)		
		CCSD(T)/CBS	MP2/CBS	CCSD(T)/CBS - MP2/CBS
1	$Fe(H_2O)$	-593.2473	-602.2960	9.05
2	$Fe(H_2O)_2$	-533.5079	-539.5072	6.00
3	$Fe(H_2O)_4$	-439.2286	-443.4756	4.25

4	$Fe(H_2O)_6$	-384.8630	-388.1550	3.29
5	$Fe(OCH_3)$	-352.0525	-354.2883	2.24
6	Fe(OOCH)	-340.2208	-343.9694	3.75
7	$Fe(SCH_3)$	-325.3746	-322.8267	-2.55
8	$Fe(C_6H_5O)$	-292.4219	-297.3581	4.94
9	$Fe(H_2O)_3(NH_2CH_3)$	-420.9803	-426.1135	5.13
10	$Fe(H_2O)_3(COOH)$	-303.6529	-307.3763	3.72
11	$Fe(H_2O)_3(OCH_3)$	-272.3165	-281.2145	8.90
12	$Fe(H_2O)_3(SCH_3)$	-263.4025	-264.0199	0.62
13	$Fe(H_2O)_5(NH_2CH_3)$	-374.0937	-426.1135	3.57
14	$Fe(H_2O)_5(COOH)$	-269.3033	-275.4854	6.18
15	$Fe(H_2O)_5(OCH_3)$	-243.2145	-252.3724	9.16
16	$Fe(H_2O)_5(SCH_3)$	-239.7555	-241.7543	2.00

Table 5 presents the CCSD(T)/CBS and MP2/CBS energies using the Helgaker extrapolation scheme as well as the difference energies value between each method. The difference between CCSD(T)/CBS and MP2/CBS energies is significant. We expect this will be happening because it means that the higher order correlation energies ($\Delta E_{CCSD(T)-MP2}$) are important for our systems. Regarding the importance of basis set with the correlation energies Jurečka *et al* [76] remarked that if the higher order correlation effects affect ordering on the energy scale more than the basis set size, which is also found in our systems, it is necessary to use at least the aug-cc-pVDZ basis set. Using bigger basis sets, for instance aug-cc-pVTZ, it is possible with more computational cost and time, nevertheless using aug-cc-pVDZ is enough for obtaining CBS energies. The extrapolation of MP2 energies to the complete basis set (CBS) clearly improves the energies, but at the same time it introduces a new small error for final values of CCSD(T) energies. The calculation of CCSD(T)/CBS using eq (35) will also introduce another small error in the final values. If we notice, the source of error will be the accumulation of errors in the calculation of MP2/CBS value and errors in the calculation of correlation energies ($\Delta E_{CCSD(T)-MP2}$). In total, these values may result in uncertainty smaller than one kilocalorie per mol. Moreover, we will use these CCSD(T)/CBS values as our reference values for benchmarking of density functional. The contributed error from the uncertainty in reference values CCSDT/CBS will slightly affect the DFT ordering performance. As such, we will not distinguish density functionals where the difference is smaller than 1 kcal/mol.

We start our discussion by comparing the $\Delta E'_{el}$ values from different systems and making a connection regarding the effect of different side chains to the systems.

Our discussion will be divided into four different small groups. Group A refers to systems of Fe complexes coordinated with only water molecules as ligands. There are four systems in this group, with an increasing number of water ligands, which are Fe with one, two, four and six water molecules. Mulliken atomic charges (Table 6.) show a consistent increasing number of the electron density of Fe(II) in $[Fe(H_2O)_4]^{2+}$ because of more charge transfer from four water ligands to Fe(II) than with $[Fe(H_2O)_2]^{2+}$. The increasing electron density in Fe(II) will consequently make the electrophilic character in Fe(II) decreases and Fe(III) becomes less prone to accept an electron and get reduced to Fe(II). With the increasing number of water ligands in Fe complexes, the value of $\Delta E'_{el}$ (Fe(II) - Fe(III)) becomes less negative and iron has a smaller tendency to get reduced. In conclusion, it is less favourable to accept one electron in the $[Fe(H_2O)_6]^{3+}$ complex than in the $[Fe(H_2O)]^{3+}$ complex. In addition to discussion of $\Delta E'_{el}$, we provide as well the bond length of complexes in Table 6.

Table 6. Bond lengths and Mulliken Atomic Charges in Group A (Fe with water ligands).

		Bond length		Mulliken Atomic Charge	
	Model	**Fe(II)-ligand (Å)**	**Fe(III)-ligand (Å)**	**Fe**	**O**
1	$Fe(H_2O)$	1.95	1.83	1.744	-0.535
2	$Fe(H_2O)_2$	1.97	1.87	1.549	-0.561
3	$Fe(H_2O)_4$	2.06	1.95	1.490	-0.606
4	$Fe(H_2O)_6$	2.16	2.04	1.734	-0.676

Fe (III) has shorter bond lengths than Fe (II). We expect that this will happen because of higher electrophilic character on Fe(III). Regarding the increasing number of water ligands with the increasing of bond length, we noticed that Fe-O bonds in $[Fe(H_2O)_6]^{2+}$ and $[Fe(H_2O)_6]^{3+}$ are the longest bonds among the four complexes. Considering the Mulliken atomic charge, Fe in $Fe(H_2O)_6$ does not have the smallest atomic charge among the four complexes. The reason of this behaviour is because Fe(II) and Fe(III) with six water ligands may be under steric tension, and this effect may also contribute to making the Fe-O bond in the hexahydrate complexes longer than the Fe-O bond in the monohydrate complexes.

Group B is composed of the Fe complexes with a single ligand that represents the side chain of one amino acid. These ligands are CH_3O^- (represents the serine (Ser) side chain), $HCOO^-$ (represents the aspartate (Asp) side chain), CH_3S^- (represents cysteine (Cys) side chain) and $C_6H_5O^-$ (represents the tyrosine (Tyr) side chain). First, we will compare ligands that contain the O atom for instance CH_3O^-, $HCOO^-$ and $C_6H_5O^-$. In metal-ligand complexes, the metal cation acts as Lewis acid and ligand acts as Lewis base. Each of these ligands has the ability for donating their lone pair

electrons to Fe. The strength of CH_3O^- , $HCOO^-$ and $C_6H_5O^-$ as Lewis base are different. Atom O in $C_6H_5O^-$ will be a stronger acid because the benzene ring is considered as electron withdrawing and stabilises the negative charge through inductive effect and also stabilises by resonance delocalization. Mulliken atomic charge of atom O in $C_6H_5O^-$ is more electronegative than atom O in CH_3O^-, therefore $C_6H_5O^-$ is a weaker Lewis base. The charge transfer from $C_6H_5O^-$ ligand to Fe is weaker compare to the charge transfer from CH_3O^- ligand to Fe. Regarding $\Delta E'_{el}$, we can conclude that it is less favourable to accept one electron for reducing $[Fe(C_6H_5O)]^{2+}$ to $[Fe(C_6H_5O)]^{+}$ than in $[Fe(OCH_3)]^{2+}$ to $[Fe(OCH_3)]^{+}$ complex. As the more nucleophilic character of the later makes it to transfer more charge to the iron, decreasing its electrophilicity and reduction propensity when compared to $[Fe(C_6H_5O)]^{2+}$complex.

Table 7. Bond lengths and Mulliken Atomic Charges in Group B (Fe-ligand X; X= CH_3O^- , $HCOO^-$, CH_3S^-, $C_6H_5O^-$)

		Bond length		Mulliken Atomic Charge	
	Model	**Fe(II)-ligand (Å)**	**Fe(III)-ligand (Å)**	**Fe**	**X (= O or S)**
1	$Fe(OCH_3)$	1.43	1.44	1.224	-0.756
2	Fe(OOCH)	1.73	1.86	1.321	-0.693
3	$Fe(C_6H_5O)$	1.88	1.77	1.086	-0.885
4	$Fe(SCH_3)$	2.17	2.39	0.848	-0.006

Table 7 shows the Fe-O bonds in different complexes with CH_3O^-, $HCOO^-$ and $C_6H_5O^-$. From the results, Fe-O bond in $C_6H_5O^-$ is the longest between three complexes. This can be connected with the Lewis base strength of the ligands. As already mentioned, the $C_6H_5O^-$ ligand is weaker base than CH_3O^- and $OOCH^-$. Therefore, it is expected the charge transfer in $C_6H_5O^-$ is weaker than in CH_3O^-, which increases the bond length. The longer bond indicates the weaker interaction between Fe-O in $[Fe(C_6H_5O)]^{+}$ complex. Furthermore, comparing CH_3O^- and $HCOO^-$, we can see that the Fe-O bond length of CH_3O^- is shorter than the one observed in $C_6H_5O^-$, and it is related with the Lewis basicity of both of them. CH_3O^- has the negative charge concentrated in a single oxygen atom, whereas $HCOO^-$ has the negative charge spreads between two equivalent oxygen atoms, making it less nucleophilic and consequently, making its bond longer. The explanation about bond lengths supports the explanation of $\Delta E'_{el}$ data. In our discussion, CH_3S^- ligand is an exception is an exception because of the difference in its S atom size. S atom has bigger size than O, and this will affect the bond length between metal and ligand.

Group C is composed of complexes with four ligands, which can adopt a tetrahedral geometry. From the structures of $[Fe(H_2O)_4]^{3+/2+}$ complexes, we changed

one of the four water molecules for one of the above-mentioned side chain-like ligands (NH_2CH_3, $HCOO^-$, CH_3O^- and CH_3S^-). First, we will compare the three different systems with Fe bound to an oxygen atom, which is $Fe(H_2O)_4$, $Fe(H_2O)_3(COOH)$ and $Fe(H_2O)_3(OCH_3)$. The calculation of $Fe(H_2O)_3(C_6H_5O)$ is not included in this study because of the convergence problem occurred on calculations using the MP2 with aug-cc-pVQZ basis set. Table 8 summarises $\Delta E'_{el}$ of all systems from Group C. The system $Fe(H_2O)_3(OCH_3)$ has the least negative energy when compared to the other systems with one atom O bound to Fe. In conclusion, the reduction of $[Fe(H_2O)_3(OCH_3)]^{2+}$ to $[Fe(H_2O)_3(OCH_3)]^{+}$ is less favourable than the reduction of $[Fe(H_2O)_3(OOCH)]^{2+}$; and the reduction of $[Fe(H_2O)_3(OOCH)]^{2+}$ is less favourable than the reduction of $[Fe(H_2O)_4]^{2+}$. The reason for this is the same as in the case of the complexes with a single ligand. If we compare with the complexes with four water ligands, we can see the similarity in $\Delta E'_{el}$ and bond length with the complex of three water ligands and one NH_2CH_3 ligand. The exchange of one water ligand for the NH_2CH_3 ligand does not give much effect in $\Delta E'_{el}$. It is because NH_2CH_3 has a neutral charge, as happens in H_2O molecules. However, it will be a different case for the ligands with a negative charge such as CH_3O^-, $HCOO^-$, CH_3S^-. There are big differences in $\Delta E'_{el}$ if we compare with complexes with four water ligands. Complexes containing the negative charge ligand will have less negative $\Delta E'_{el}$.

Table 8. $\Delta E'_{el}$ (Fe(II) - Fe(III)) CCSD(T)/CBS energy of of Group C.

	Model	$\Delta E'_{el}$ (Fe(II) - Fe(III)) (kcal/mol) CCSD(T)/CBS
1	$Fe(H_2O)_4$	-439.2286
2	$Fe(H_2O)_3NH_2CH_3$	-420.9803
3	$Fe(H_2O)_3(OOCH)$	-303.6529
4	$Fe(H_2O)_3(OCH_3)$	-272.3165
5	$Fe(H_2O)_3SCH_3$	-263.4025

Regarding the bond lengths, the size of both atoms N and O is very similar. Therefore, even though the Fe-N bond length in $Fe(H_2O)_3NH_2CH_3$ is longer than the Fe-O bond length in $Fe(H_2O)_4$, but the difference is only 0.03 Å. In another case, the Fe-N bond length in complex $Fe(H_2O)_3NH_2CH_3$ is longer than the Fe-O bond length in $Fe(H_2O)_3(OCH_3)$. This happens because the charge transfer in Fe-O is stronger than in Fe-N, and because the oxygen atom can donate more lone pairs electron to Fe than the nitrogen atom. Similar to Group B, the Fe-O bond length of CH_3O^- in $Fe(H_2O)_3(OCH_3)$ is shorter than Fe-O bond length of $HCOO^-$ in $Fe(H_2O)_3(COOH)$. The reason is similar to the previous system and it is related with the Lewis basicity of both of them.

Table 9. Bond lengths in Group C ([Fe-$(H_2O)_3$(ligand X)]; (X= NH_2CH_3, H_2O, $HCOO^-$, OCH_3^- and CH_3S^-).

	Model	Bond length Fe-ligand (Å)	Bond length Fe-ligand (Å)
1	$Fe(H_2O)_3NH_2CH_3$	2.09	2.01
2	$Fe(H_2O)_4$	2.06	1.95
3	$Fe(H_2O)_3(COOH)$	1.91	1.76
4	$Fe(H_2O)_3(OCH_3)$	1.79	1.69
5	$Fe(H_2O)_{3(}SCH_{3)}$	2.23	2.43

Group D is composed of complexes with six ligands that adopt an octahedral geometry. From $Fe(H_2O)_6$ complex, we changed one water ligand by one of the other different side chain-like ligands such as NH_2CH_3, $HCOO^-$, CH_3O^- and CH_3S^-. The complexes with octahedral geometry show a similar behaviour to the complexes with tetrahedral geometry. The complex of $Fe(H_2O)_5(NH_2CH_3)$ has the most favourable $\Delta E'_{el}$ for Fe(III)/Fe(II) reduction among the complexes with five water ligands and one amino acid side chain ligand. In the same magnitude with Group C, Group D follows the trend where $\Delta E'_{el}$ of $Fe(H_2O)_6$ and $\Delta E'_{el}$ of $Fe(H_2O)_5(NH_2CH_3)$ do not differ significantly. The exchange of one H_2O ligand by NH_2CH_3 does not give much effect in $\Delta E'_{el}$ because both of them are neutral charged ligands. Furthermore, the comparison between complexes that contain an N atom in $Fe(H_2O)_5(NH_2CH_3)$ and an O atom in $Fe(H_2O)_5(OCH_3)$ shows that the $Fe(H_2O)_5(OCH_3)$ complex appears to have less negative energy. Moreover, $Fe(H_2O)_5(OCH_3)$ complex has the least negative energy among the three systems with an oxygen atom bound to Fe. From both previous facts, it can be concluded that $Fe(H_2O)_5(OCH_3)$ is the least favourable for reduction of $[Fe(H_2O)_5(OCH_3)]^{2+}$ to $[Fe(H_2O)_5(OCH_3)]^+$; excluding $Fe(H_2O)_5(SCH_3)$ complex. In conclusion, the ranking of the tendency for the reduction of the complexes with different types of ligands is the following: $H_2O > NH_2CH_3 > HCOO^- > CH_3O^- > CH_3S^-$

Table 10. $\Delta E'_{el}$ (Fe(II) - Fe(III)) CCSD(T)/CBS energy of Group D.

	Model	$\Delta E'_{el}$ (Fe(II) - Fe(III)) (kcal/mol) CCSD(T)/CBS
1	$Fe(H_2O)_6$	-384.8630
2	$Fe(H_2O)_5(NH_2CH_3)$	-374.0937
3	$Fe(H_2O)_5(COOH)$	-269.3033
4	$Fe(H_2O)_5(OCH_3)$	-243.2145
5	$Fe(H_2O)_5(SCH_3)$	-239.7555

Moreover, we analysed the bond lengths between Fe-N and Fe-O in the complexes from Group D (data showed in Table 11). Group D has a similar behaviour to Group C. Fe-N bond length in complex $Fe(H_2O)_3NH_2CH_3$ is longer than Fe-O bond

length in $Fe(H_2O)_3(OCH_3)$. It is because of the bigger charge transfer from O atom in CH_3O^- to Fe cation than atom N in NH_2CH_3. The oxygen atom can donate more lone pairs electron to Fe than the nitrogen atom. We can also compare the results between the complexes in which the interaction to iron occurs by an oxygen atom. It reveals that Fe-O in $Fe(H_2O)_5(OCH_3)$ has the shortest bond length among the three other complexes. The shortest bond length can give an indication of the strength of charge transfer from ligand to metal. It means charge transfer from ligand atom O of CH_3O^- ligand to Fe is stronger than the one occurred in the other complexes. In this Group and previous Groups, the size of atom S invalidates the direct comparison of the bond length of SCH_3 with other ligands.

Table 11. Bond lengths in Group D ([Fe-$(H_2O)_5$-ligand X)]; (X= NH_2CH_3, H_2O, $HCOO^-$, CH_3O^- and CH_3S^-).

	Model	Bond length Fe-ligand (Å)	Bond length Fe-ligand (Å)
1	$Fe(H_2O)_5(NH_2CH_3)$	2.17	2.08
2	$Fe(H_2O)_6$	2.16	2.04
3	$Fe(H_2O)_5(COOH)$	2.08	1.84
4	$Fe(H_2O)_5(OCH_3)$	1.89	1.71
5	$Fe(H_2O)_5(SCH_3)$	2.33	2.51

4.3 Benchmarking of DFT functional

We have benchmarked 45 density functionals which consisted of 2 LDA, 7 GGA, 6 M-GGA, 10 H-GGA, 10 HM-GGA, 4 double H-GGA, 1 range-separated GGA, 1 NGA, 1 range-separated HM-GGA, 1 range separated HM-NGA, 1 M-NGA and 1 GGE. $\Delta E'_{el}$ is obtained from the subtraction of single point energy of Fe (II) and Fe (III) like it were described by eq (14). The benchmarking study of DFT functionals was performed using the basis set 6-311++g(2df,2p). We have chosen a large basis set in order to minimise the basis set truncation error, making that the performance comes essentially from the DFT and not from the basis set. This study does not represent the global quality of the functionals, which should be measured for different chemical properties for all systems. In particular, we focused on evaluating the performance of various density functionals for predicting the $\Delta E'_{el}$ for different metal complexes with Fe and different amino acid side chains. The determination of $\Delta E'_{el}$ will be essential for the study of electron affinity properties, which are fundamental to calculate the reduction potentials.

In the same manner with the previous discussion, we will compare the performance of functionals for describing $\Delta E'_{el}$ values from different systems in comparison with the CCSD(T)/CBS $\Delta E'_{el}$ values. Our aim is to find which functionals give better performance for the four small groups (A, B, C and D) of systems and for all different Fe-ligands systems. First, we will compare the performance of functionals for Group A (systems of Fe coordinated with only water molecules). Table 12 summarises the mean unsigned error (MUE of $\Delta E'_{el}$) and maximum error (MaxE of $\Delta E'_{el}$) between single point DFT/6-311++g(2df,2p) energies with CCSD(T)/CBS energies. We only show the values for the ten functionals with the best performance for Group A. The complete performance for all density functionals can be found in supporting information (SI).

Table 12. Ten density functionals with the best performance for Group A complexes.

	Functional	**Type**	**%HF$_{exchange}$**	**Mean MUE**	**MaxE**	**MUE**			
						Fe (H_2O)	**Fe $(H_2O)_2$**	**Fe $(H_2O)_4$**	**Fe $(H_2O)_6$**
1	MPWB1K	hm-GGA	44	0.61	1.06	0.98	0.02	1.06	0.38
2	BB1K	hm-GGA	42	0.74	1.61	0.12	1.05	1.61	0.20
3	MPW1N	h-GGA	40.6	1.34	1.93	1.79	1.29	0.33	1.93
4	MPW1K	h-GGA	42.80	1.73	2.27	2.27	2.04	0.64	1.95
5	M06-2X	hm-GGA	54	2.17	5.08	5.08	2.53	0.45	0.63
6	BMK	hm-GGA	42	2.18	3.97	3.97	0.26	2.44	2.06
7	MN12-SX	hm-NGA	25	2.71	5.87	1.37	2.69	0.90	5.87
8	MPW1B95	hm-GGA	31	2.95	5.37	1.84	5.37	3.90	0.66
9	B3LYP	h-GGA	20	3.15	6.67	0.10	6.67	3.96	1.86
10	B3PW91	h-GGA	20	3.52	7.51	1.73	7.51	4.69	0.15

The values above are important for understanding the accuracy of each functional to describe the reduction potential in metalloproteins or biological metal systems, in which the EA is one of the most important properties connected with it. We have divided the performance of density functionals into three groups according to the MUE of $\Delta E'_{el}$. In group I, the MUE values range from 0 to 2.31 kcal/mol (which corresponds to an error below 0.1 V for 1 electron transfer). Group II contains functionals with MUE between 2.31 and 4.62 kcal/mol (which corresponds to an error between 0.1 and 0.2 V for 1 electron transfer). Whereas, Group III is made up of functionals with MUE more than 4.62 kcal/mol. Table 12 summarises the ten density functionals, which give the best performance for Group A (systems of Fe complexes with only water ligands). We coloured those functionals that give MUE values lesser

than 2.31 kcal/mol with grey shade in Table 12 and for the rest of discussion as well. The density functionals that give MUE and MaxE values lesser than 2.31 kcal/mol are MPWB1K, BB1K, MPW1N, MPW1K, M06-2X and BMK. We also remark that the reference is not accurate enough to distinguish which is the best one within the first four functionals because they only differ by ~1 kcal/mol or less. All the functionals that have good performances are non-local functionals and come from hybrid meta-GGA and hybrid GGA. If we notice, the first four functionals have high HF exchange percentages (with values more than 40%). The very popular B3LYP functional ranks 9th in the terms of MUE and 11th in the terms of MaxE, belonging to the Group II.

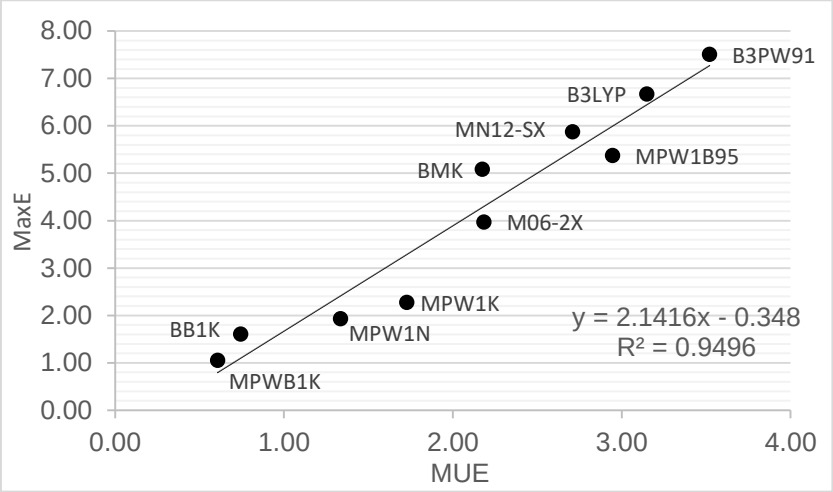


Fig 5. MUE and MaxE values of the best 10 density functionals for systems of Group A.

In addition, we plot the graphic to see the correlation of MUE and MaxE from these ten functionals. Both MUE and MaxE values are important for assessing the performance of density functionals. MUE is the average unsigned error given by functionals for four different systems. Furthermore, MaxE can give us a prediction if we have a particular case, in which the functionals fail to give a better description in the four systems. From the Figure 5, we can see that in general the MUE and MaxE values will increase linearly. When the MUE increases, the value of MaxE will also increase. MUE and MaxE values give results that support each other, and we can say that the MUE and MaxE values are reliable for describing the performance of density functionals for Group A.

Second, we compare the performance of functionals for Group B (systems of Fe with a single different side chain ligand). Table 13 summarises the mean unsigned error (MUE of $\Delta E'_{el}$) and the maximum error (MaxE of $\Delta E'_{el}$) between single point DFT/6-311++g(2df,2p) energies with CCSD(T)/CBS energy values of ten functionals with the best performances for Group B. We remarked that the functionals that work

better for Group B are mostly different from those that work better for Group A. The density functional that gives the MUE value less than 2.31 kcal/mol are PBE1PBE, MPW1B95 and B1B95 which make them belong to Group I. There are eight functionals included in Group II with MUE less than 4.62 kcal/mol. The well-known B3LYP function ranks 9th in the term of MUE (Group II) and 6th in the term of MaxE (Group III). Further, we observe that the three density functionals which give better performance are the hybrid GGA and hybrid meta-GGA functional. Therefore, we can conclude that the functionals which give a better performance for System B are hybrid GGA and hybrid meta-GGA functionals.

Table 13. Ten density functionals with the best performance for Group B complexes.

	Functional	**Type**	**%HF$_{exchange}$**	**Mean MUE**	**MaxE**	**MUE**			
						Fe (OCH_3)	**Fe (OOCH)**	**Fe (C_6H_5O)**	**Fe (SCH_3)**
1	PBE1PBE	h-GGA	21	0.94	1.88	0.33	-0.20	1.36	-1.88
2	MPW1B95	hm-GGA	31	1.40	3.27	0.11	0.38	-1.83	-3.27
3	B1B95	hm-GGA	28	1.81	3.53	1.19	-3.53	-0.22	2.30
4	B1LYP	h-GGA	25	2.54	7.72	0.76	0.91	7.72	0.76
5	B3PW91	h-GGA	20	2.62	5.11	-0.10	-4.56	0.72	-5.11
6	BMK	hm-GGA	42	3.02	7.04	-7.04	1.93	3.08	0.03
7	wB97X-D	h-GGA	22.20/100	3.12	5.34	-3.05	-0.56	3.52	-5.34
8	MPWB1K	hm-GGA	44	3.63	6.96	1.02	6.96	5.05	1.49
9	B3LYP	h-GGA	20	3.98	6.41	-1.97	-6.41	2.06	-5.49
10	BB1K	hm-GGA	42	4.33	6.82	-3.46	6.82	5.52	1.50

Figure 6 shows the correlation of MUE and MaxE from these ten functionals. MUE gives an idea about the average accuracy performance of density functional in the four different systems of Group B. There is an interesting fact, in which we observed that the MaxE for each density functional is not always provided by the same ligand. From Table 13, we remark that complexes with the ligands CH_3S^- and $HCOO^-$ share the same source as the main contributor for MaxE value of the system. The only system that has a less contribution of MAE for Group B is the complex with CH_3O^-. Comparing with Group A, the MUE and MAE values of Group B have less linearity. We can explain this behaviour by comparing B1LYP and B3PW91 as an example. In the term of MUE value, B1LYP has 2.54 kcal/mol, which is lower than B3PW91. In contrary, in the term of MaxE value, B1LYP has 7.72 kcal/mol that is higher than B3PW91. In particular, B1LYP has smaller values of MUE for the systems with CH_3O^- (MUE = 0.76 kcal/mol), $HCOO^-$ (MUE = 0.91 kcal/mol) and CH_3S^- (MUE = 0.76 kcal/mol), however, there is a dramatic increase in the MUE for the system with $C_6H_5O^-$

(MUE = 7.72 kcal/mol). In one case when one functional works accurately for some systems and decrease suddenly for one particular system, this will have an impact on the overall accuracy of the density functional. Therefore, for assessing the accuracy of functionals, we have to consider both MUE and MaxE values. In Figure 6, if we exclude MUE and MaxE value from PBE1PBE, the linearity will decrease (value of R^2 = 0.5146).

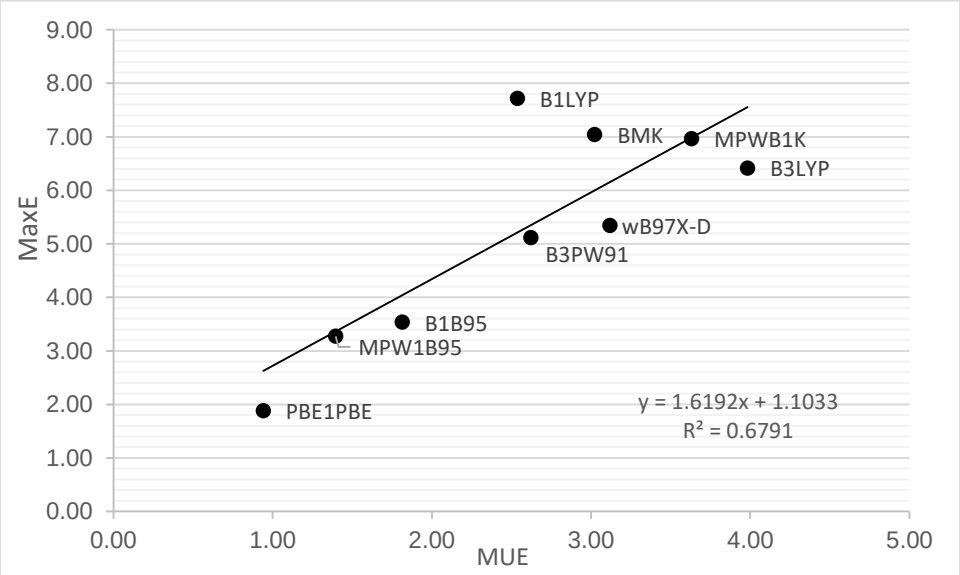


Fig 6. MUE and MaxE values of the best 10 density functionals for systems of Group B.

Thirdly, we analyse the DFT performance on systems of Group C that consist of complexes with four ligands bound to iron (three water molecules plus one amino acid side chain). Table 14 displays the mean unsigned error (MUE of $\Delta E'_{el}$) and maximum error (MaxE of $\Delta E'_{el}$) values between single point DFT/6-311++g(2df,2p) energies and CCSD(T)/CBS energetic values of ten functionals with the best performance for Group C. We note the functionals which work better for System C are mostly different with System B yet similar with System A. The density functionals which give MUE values lesser than 2.31 kcal/mol are BMK, MPW1N, M06-2X, MPW1K, MN12-SX, MPWB1K and MN12-L. All of these functionals constitute the Group I.

Table 14. Ten density functionals with the best performance for Group C complexes.

	Functional	**Type**	**%HF$_{exchange}$**	**Mean MUE**	**MaxE**	**MUE**			
						Fe(H$_2$O)$_3$ (NH$_2$CH$_3$)	**Fe(H$_2$O)$_3$ (COOH)**	**Fe(H$_2$O)$_3$ (OCH$_3$)**	**Fe(H$_2$O)$_3$ (SCH$_3$)**
1	BMK	hm-GGA	42	1.14	2.15	-0.66	2.15	-0.20	1.54
2	MPW1N	h-GGA	40.6	1.76	4.33	-1.41	1.13	-0.18	4.33
3	M06-2X	hm-GGA	54	1.87	3.43	-1.03	0.22	-3.43	2.81
4	MPW1K	h-GGA	42.80	1.93	4.60	-1.83	0.55	-0.71	4.60

5	MN12-SX	hm-NGA	25	2.03	3.46	-2.84	-0.28	-3.46	-1.56
6	MPWB1K	hm-GGA	44	2.09	4.27	-0.43	2.78	0.90	4.27
7	MN12-L	hm-NGA	0	2.26	4.21	0.62	3.69	-0.53	-4.21
8	BB1K	hm-GGA	42	2.64	4.63	0.25	3.78	1.92	4.63
9	M05-2X	hm-GGA	56	3.37	4.82	-3.43	-2.20	-4.82	3.05
10	B3LYP	h-GGA	20	3.55	7.07	2.73	7.07	3.49	0.93

The functionals in Group I are mainly hybrid meta-GGA and hybrid GGA. In particular, most of this hybrid meta-GGA functionals have HF exchange more than 40%. Further analysis of MUE for each system can be used to find which system has a bigger MaxE than the others. The system with $(H_2O)_3(SCH_3)$ is the main contributor of the MaxE values in System C. The correlation of MUE and MaxE is enclosed in Figure 7. There is a particular behaviour which functional MPW1K has a lower value than MN12-SX in term of MUE but higher value than MN12-SX regarding MaxE value. This behaviour affects the linearity of the ten functionals when we plot them in one graphic in Figure 7.

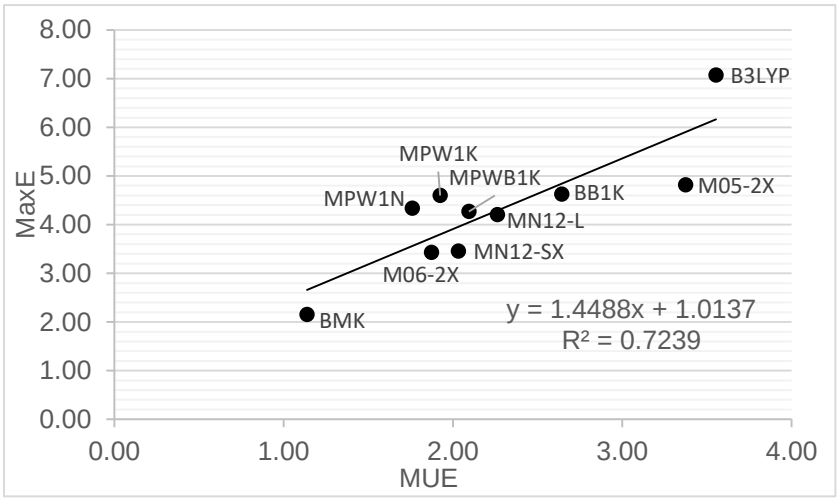


Fig. 7 MUE and MaxE values of the best 10 density functionals for systems of Group C.

The Group D consists of complexes with six ligands coordinated to iron. There are five water molecules and one different amino acid side chain for each system. Table 15 presents the mean unsigned error (MUE of $\Delta E'_{el}$) and maximum error (MaxE of $\Delta E'_{el}$) value of ten functionals with the best performances for Group D systems. There are nine functionals which have MUE value less than 2.31 kcal/mol. The functionals are wB97X-D, MPWB1K, B3PW91, BB1K, MPW1B95, M06-2X, B3LYP, BMK, and MN12-L.

Table 15. Ten density functionals with the best performance for Group D complexes.

	Functional	Type	%$HF_{exchange}$	Mean MUE	MaxE	MUE			
						$Fe(H_2O)_5(NH_2CH_3)$	$Fe(H_2O)_5(COOH)$	$Fe(H_2O)_5(OCH_3)$	$Fe(H_2O)_5(SCH_3)$
1	wB97X-D	h-GGA	22.20/100	1.32	2.13	-2.06	-2.13	0.70	-0.39
2	MPWB1K	hm-GGA	44	1.47	4.36	-0.55	-0.94	0.03	4.36
3	B3PW91	h-GGA	20	1.63	3.15	0.14	0.24	3.15	2.98
4	BB1K	hm-GGA	42	1.63	4.85	-0.32	-0.50	0.86	4.85
5	MPW1B95	hm-GGA	31	1.87	3.63	0.51	0.52	2.81	3.63
6	M06-2X	hm-GGA	54	2.08	3.69	-0.02	-1.79	-3.69	2.81
7	B3LYP	h-GGA	20	2.09	3.45	-1.85	-1.56	1.49	3.45
8	BMK	hm-GGA	42	2.21	3.18	-1.76	-2.67	-1.24	3.18
9	MN12-L	hm-NGA	0	2.21	3.97	-3.97	-2.85	-0.71	1.31
10	MPW1N	h-GGA	40.6	2.50	4.40	-1.75	-2.44	-1.41	4.40

We noticed that the density functionals which give the best performances for the systems on Group D are mainly hybrid GGA and hybrid meta-GGA. We tried to find out which system is the main contributor to MaxE value. Within these ten different functionals, the $Fe(H_2O)_5(OSH_3)$ system is the main contributor to the large average value of MaxE. Furthermore, we analysed the correlation of MUE and MaxE, which is showed in Figure 8. We also find a similar case as occurred in Group C, in which one functional has lower MUE value but a bigger MaxE value than other functional, for instance, MPWB1K and B3PW91. This behaviour affects the linearity of the ten functionals.

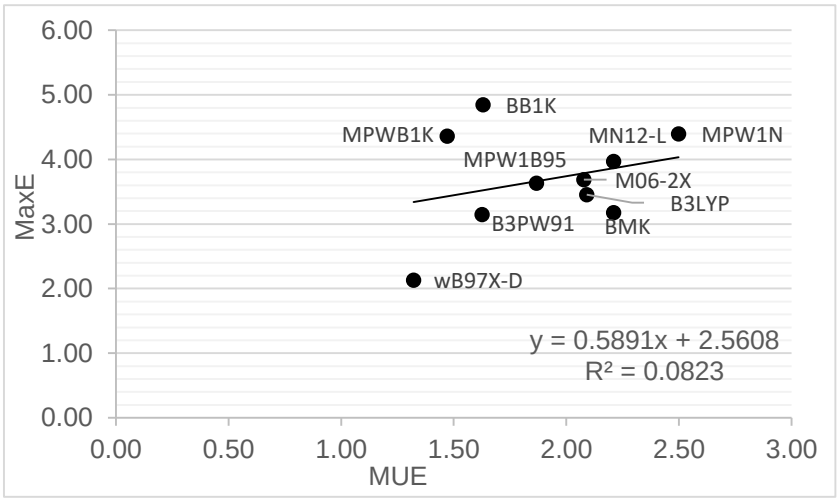


Fig 8. MUE and MaxE values of the best 10 density functionals for systems of Group D.

Both of MUE and MaxE values are important for assessing the performance of density functionals. MUE is the average of errors given by functionals for the four

different systems, whereas MaxE can give us a prediction if we have a particular case, in which a functional fail to give the best description of the four systems included in this Group. When the increment of MUE values is linear with the increasing of MaxE values, it means that both of values are reliable and we can consider one of the values for assessing the performance of functionals. However, in some particular cases, we find that suddenly the MaxE value increases drastically. In this case, it means the functional works well for the most systems, however, fails drastically for one particular system. For example, in the case of functional MPWB1K and BB1K, the functional works well for system $Fe(H_2O)_5(COOH)$, $Fe(H_2O)_5(OCH_3)$, $Fe(H_2O)_5(NH_2CH_3)$ and does not work well for $Fe(H_2O)_5(SCH_3)$ system.

Table 16. Ten density functionals with the best performance for all Fe-ligands systems.

	Functional	**Type**	**%$HF_{exchange}$**	**MUE**	**MaxE**
1	MPWB1K	hm-GGA	44	1.95	4.16
2	BMK	hm-GGA	42	2.14	4.08
3	BB1K	hm-GGA	42	2.34	4.47
4	MPW1B95	hm-GGA	31	2.60	4.79
5	MPW1N	h-GGA	40.6	2.79	4.54
6	M06-2X	hm-GGA	54	2.94	6.00
7	B3PW91	h-GGA	20	3.08	5.96
8	B3LYP	h-GGA	20	3.19	5.90
9	MPW1K	h-GGA	42.80	3.25	5.01
10	PBE1PBE	h-GGA	21	3.39	5.33

Further discussion is about the performance of functionals in characterising the $\Delta E'_{el}$ values for all sixteen model systems. Table 16 summarises the performance of the ten best functional for all systems. The functional which include in Group I (MUE value less than 2.31 kcal/mol) are MPWB1K and BMK. They are closely followed by BB1K, MPW1B95, MPW1N and M06-2X functionals. The popular functional B3LYP ranks in the 8th in term of MUE value. Even though MPWB1K come as the first in the term of MUE, it is hard to compare which functional has better performance than others which lie between BMK until B3LYP. They have the comparable difference of ~1kcal/mol, which make their performance very difficult to distinguish. There is interesting fact from these ten functionals in which mostly of them are hybrid meta-GGA and hybrid GGA functionals. We can conclude that the hybrid meta-GGA with higher HF exchange percentage and hybrid GGA work better for our 16 Fe-ligands systems.

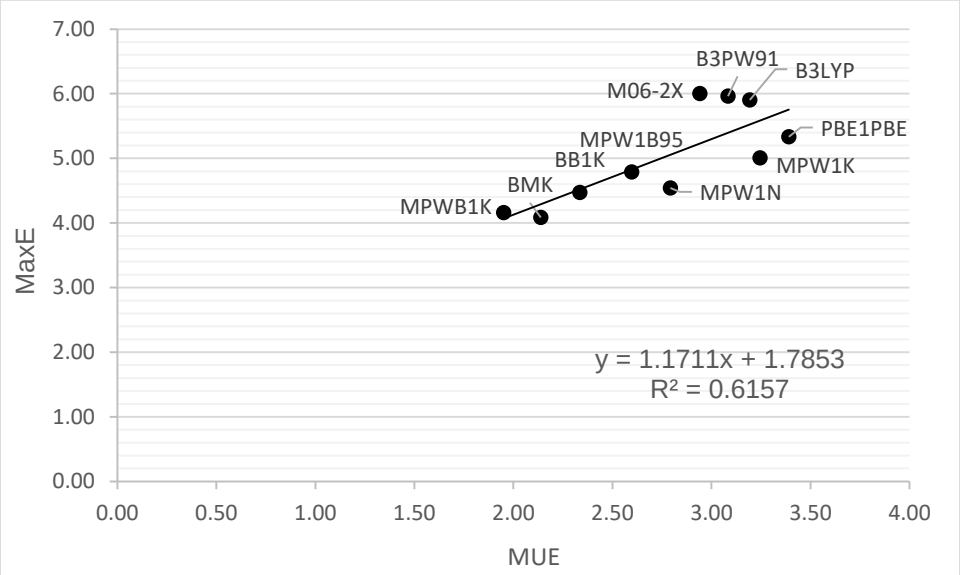


Fig.9 MUE and MaxE values of the best 10 density functionals for all 16 systems.

In addition, we analyse the relation of MUE and MaxE for all systems.There are some cases in which one functional has bigger value in terms of MUE but the bigger value in terms of MaxE. B3PW91 has less MUE value than PBE1PBE, however, the bigger value in terms of MaxE. However, we found that there is no sudden increase drastically in MaxE value for each responsible MUE. The values of MUE and MaxE are reliable as parameters for assessing the performance of functionals.

Zhao and Truhlar [77] have conducted studies on density functionals for transition metal complexes. In their studies, they have concluded that M05-2X and M06 functionals are suitable functionals for predicting the bond dissociation energies of phosphine binding energies of ruthenium precatalysts for olefin metathesis. M062-X has a high percentage of HF exchange and stated does not perform well for describing transitional metal chemistry. However, in one of our results we found that M06-2X performs well in our systems and give the MUE value of 2.94 kcal/mol for the total systems. In our model systems, we have used iron as transition metal and we have calculated the $\Delta E'_{el}$ as the main contribution of EA in the system. We cannot directly compare our studies with Zhao and Truhlar studies because used a different transition metal cation and our objectives for these studies are different

Luo *et al.* [78] have studied the density of functional theory of 3d series transition metal atoms (Sc to Zn) and their cation. They have analysed the energy of the open shell states using a new broken-symmetry method called reinterpreted broken symmetry (RBS) method. They have calculated the ionization potential calculation with RBS method and compared them with experimental data. They have found out that

wB97X-D functional have the MUE of < 5 kcal/mol. They have recommended the other functionals such as MPW1B95 and MPWB1K. In line with our current work, we have found out that wB97X-D functional perform well for complexes with six ligands coordinated to iron with the smallest MUE (1.32 kcal/mol) among all functionals tested in the current work. In addition, we have also found out that MPWB1K (MUE = 1.95 kcal/mol) and MPW1B95 (MUE = 2.60 kcal/mol) perform well for our 16 Fe-ligands systems.

Riley and Merz [11] have studied the performance of several density functionals to compute the heat of formation and ionization potential (IP) for systems containing third-row transition metals. They have concluded that the inclusion of exact exchange term in DFT methods gives more accurate results of heats of formation and IP of transition metals systems. The hybrid GGA functional, B3LYP, was found to produce the lowest overall error for IP, when combined with 6-31G** and TZVP basis set, respectively. It has 19.6 kcal/mol and 11.07 kcal/mol of errors in comparison with the experimental results. Yang *et al*. [79] continue the study of Riley and Merz [11] and had studied the IP values using DFT and effective core potential (ECP) methods for transition metal complexes from Ti to Zn. They analysed various systems of transition metals with a binding partner, for instance, H, N, O, S, F, Cl, OH, CO and CH_3 ligands, and they compare their results with experimental data. They remarked that the functionals with the best performance for iron are TPSS1KCIS and B3LYP, with MUE values of 5.76 and 5.99 kcal/mol respectively. Moreover, they have concluded that there is no obvious tendency for the IP assessment: on the one hand, the hybrid GGA functionals are better than the all other classes; while on the other hand, the hybrid meta-GGA functionals become more disappointing in performance. They have concluded that in general, the B3LYP gives the best performance for all the transitional metal systems, with the exception of Ti complexes. Our results show that the ten functionals which have the best performance for all 16 systems are mostly hybrid GGA and hybrid meta-GGA They perform quite well for describing $\Delta E'_{el}$, which is essential for EA properties and closely related with reduction potential studies.

This present study is an important first step for further studies involving the transitional metal system. To begin with, it shows that BMK is the most accurate functional for describing the $\Delta E'_{el}$ in the Fe(III)/Fe(II) system coordinated with water and amino acid side chains. In general, B3LYP is still quite accurate among the ten best functionals to predict $\Delta E'_{el}$. In addition, if we want to be more specific, for example determining EA in a system of iron that contains many water ligands, it will be better to

use MPWB1K, BB1K, MPW1N and MPW1K which had shown more accurate results for these particular systems. Similarly, for the systems containing methoxy, carboxylic acid or thiol, the PBE1PBE is the best one for describing the EA. It is important to note that redox reactions occur in solution phase. However, the results provided by this work show only how well DFT methods describe the gas-phase $\Delta E'_{el}$ that is crucial for characterise EA properties.

5. Conclusion

The results presented in this work are important for the characterization of EA properties, which are key properties for further studies of redox reactions in biological systems. A series of benchmarking data of $\Delta E'_{el}$ for EA properties is presented. $\Delta E'_{el}$ for EA reference values were determined at the very accurate CCSD(T)/CBS level of theory. These values were obtained from the sum of the complete basis set limit of MP2 energies and the CCSD(T) correction term evaluated using the aug-cc-pVDZ basis set. The MP2/CBS energies were determined using the Truhlar and Helgaker extrapolation schemes. The difference between E'_{el} values at MP2/CBS level extrapolated with both schemes is very small (absolute average is 0.31 kcal/mol). The results from DFT benchmarking conclude that for all systems studied, both hybrid meta-GGA and hybrid GGA density functionals work well. MPW1K and BMK (MUE 1.95 and 2.14 kcal/mol) functionals are the ones with the best performance for the sixteen Fe-ligands complexes.

6. Further Work

From our work, we have noticed that the MUE value increases with the presence of amino acid side chains as described in Group B, Group C and Group D. In reality, most of the irons in protein bound with at least four amino acid side chains. Therefore, further work will involve, for consideration, the coordination of shells which are more similar to the coordination environment in proteins, with one water and the rest being Asp/Glu/Cys/Tyr/Ser amino acid side chains.

II. Impact of Polarity of Fullerene Derivatives on the Morphology of Bulk Heterojunction Solar Cells from Solvent Evaporation Simulations

1. Introduction

Organic photovoltaics (OPV) is one of the interesting area studies for the experimental and theoretical researcher because of their several advantages as compared to inorganic photovoltaics which is environmentally friendly, low-cost production and suitable for flexible devices [80, 81]. OPV consists of organic materials for instance conjugated polymer which has p or n-type semiconducting properties [82]. The p-type semiconductor refers to the positive charge of the hole whereas n-type semiconductor refers to the negative charge of the electron. The low-cost production of OPV is connected to the prospect of using printing or coating from solution [83]. The possibility of using low-cost manufacture process lead to the growing interest for large scales production of OPV. In the last decade, there are a lot of studies regarding organic material which generates better photovoltaic energy conversion [84, 85].

The history of OPV start with the application of organic semiconductor material to directly replace the inorganic semiconductor material in solar cell and the reported performance was very poor [83]. Hiramoto *et al.* [86, 87] introduced the concept of an organic cell structure by stacking two heterojunction devices. From a series of studies, it had been reported the increasing of photoconductivity upon the C_{60} addition to the conjugated polymer. This information leads to the development of the polymer-fullerene bilayer heterojunction and bulk heterojunction (BHJ). The BHJ concept is similar to the concept to the Hiramoto. BHJ was introduced by Yu *et al* [88, 89] that blended a polymer and a fullerene derivative together, two materials having donor and acceptor electron material properties. One of the most employed donor materials is polymer poly(3-hexylthiophene) (P3HT) whereas the electron acceptor material is [6,6]-phenyl-C61-butyric acid methyl ester (PCBM) [90].

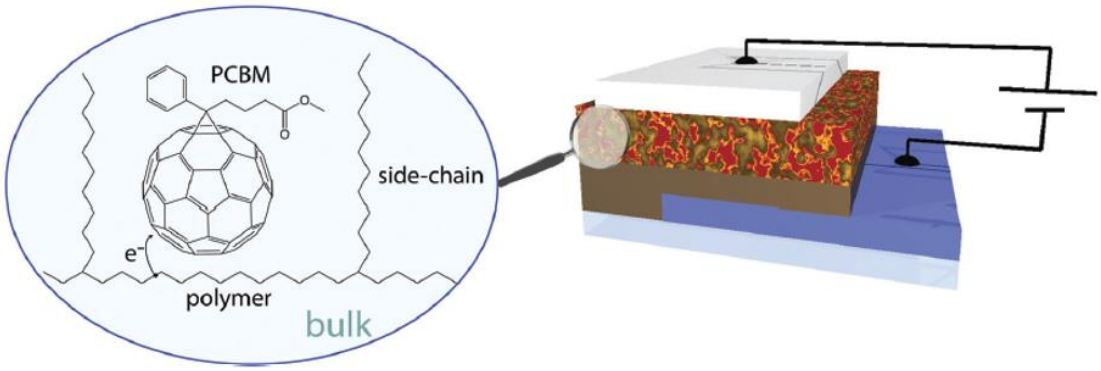


Fig. 10 Bulk Heterojunction (BHJ) solar cell. Picture was taken from de Gier *et al*. [91]

One of the main essential difference between OPV and inorganic photovoltaic (IPV) is the photoconversion mechanism. The light absorption in OPV leads to the production of exciton whereas in IPV leads to the generation of free electron-hole pairs [92]. In general, the basic working principle of BHJ OPV can be explained into five significant steps described in Figure 11. The first step involving the light absorption and exciton generation in the active layer. The second step is exciton diffusion in which the generated exciton has to diffuse to donor-acceptor interface where the charge separation can take place. The third step is exciton dissociation to form the free charges at the donor-acceptor interface. The fourth step is charge separation. The charge separation involves an intermediate charge transfer state on the donor-acceptor interface. The last step is the charge transport in which the electron and the hole move in opposite directions toward the two electrodes and collection of the charges [93].

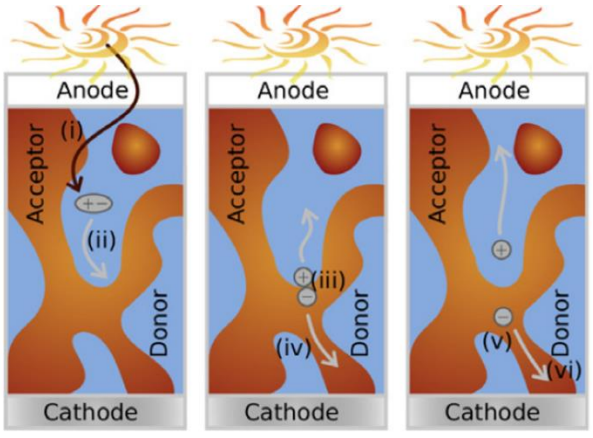


Fig. 11 Processes which happen in a BHJ. Picture was taken from Wodo *et al*. [93]

The development of studies and experiments regarding OPV is still growing until now. The main goal of the studies is increasing the power conversion efficiencies of OPV. These experiments lead to the discovery of the possible candidate for substituting the PCBM and P3HT. At this moment, one of the most promising

candidates is thieno[3,4-b]thiophene-altlbenzodithiophene [PTB7] and $PC_{70}BM$ [85, 94]. The modification of fullerene derivatives has been one of the options since completely replace PCBM with another material has not been successful until now [81]. The increasing number of research for developing better acceptor material reveals another challenging problem regarding the morphology studies of synthesised solar cell. The importance of morphology mainly affects the performance of a solar cell. Several studies had mentioned in which the good morphologies result in better and more efficient solar cell than the one with bad morphologies [95]. However, the definition of good and bad morphologies are still hard to categorise using quantitative measurement. Understanding the morphology's role in the performance of the device is an essential step for completely achieve the most desirable morphology.

Recently, Matsumoto *et. al.* [81] have successfully synthesised new fullerene derivatives with different polarity. One of the fullerene derivatives with high polarity has been named DMPCBM whereas one of the low polarity ones has been named DMPCHp. Their experimental works show fullerene derivatives with lower polarity has a better power of conversion of efficiency (PCE) than higher polarity. In addition, the high polarity fullerene derivatives do not mix as good with P3HT due to a lower affinity caused by the polarity mismatch. Evaluation of phase separation of P3HT with fullerene derivatives can be studied using the morphological characterization. The final morphological structure of BHJ thin film is essential for studying the efficiency of new fullerenes derivatives. Interaction of new fullerenes derivatives with polymer P3HT also plays an essential role for final morphology appearance obtained in thin films. Morphological analysis can be used for finding the compatibility between new fullerenes derivatives and P3HT. However, further computational studies about the morphology of these two new fullerenes derivatives have not been performed.

In this study, we attempted to investigate the effect of the polarity of fullerenes derivatives on the morphology using solvent evaporation simulations and develop the coarse-grained (CG) simulation for fullerenes derivatives. Our main priority is to test whether our simulation will be in agreement results with the experimental findings by Matsumoto *et al.* [81].

2. Models and Methods

2.1 The ideal morphology for BHJ solar cells

In the introduction, we have talked about the importance of morphology for the performance of the solar cells. The morphology of BHJ the active layer is the outcome of the interaction between the electron donor and electron acceptor materials. The ideal nanomorphology of the active layer for efficient BHJ requires the formation of continuous interpenetrating networks and separated donor and acceptor phases. The exciton generated in any spot of the active layer can diffuse to donor and acceptor interface to dissociate into free electrons and holes which can be transported to the respective electrodes before recombining with each other.

(a) (b)

Fig. 12 (a) Typical morphology BHJ (b) ideal morphology BHJ. Picture was taken from Wodo *et al* [93]

In Figure 12 (a) which has a more random pattern is the typical morphology obtained from the experiment in the laboratory. In the right side of Figure 12 (b), which has more ordered pattern is the ideal morphology of BHJ. In general, it is hard to obtain the ideal morphology of BHJ. However, creating the short and continuous pathways of donor electron and acceptor electron regions which connected to the electrodes is essential since every step in BHJ process in affected by the morphology of the thin film.

2.2 Molecular Dynamics simulation

Computational techniques regarding the morphologies of BHJ involve in molecular dynamics (MD) simulation. MD simulation can help us to understand the molecular interaction happened in BHJ solar cells. MD simulation is a way to generate the equilibrium ensemble for the system. The classic simulation depends on the empirical approximation called force field for calculating the interaction and evaluate the potential energy of the system [96]. A force field contains a set of equations for calculating the potential energy and force from particle coordinates including the parameter used in the equations. The force fields usually consist of two potential functions. Nonbonded interactions consist of the Lennard-Jones repulsion and dispersion as well as Coulomb electrostatics. Bonded interactions account covalent

bond-stretching, angle-bending, torsion potentials and out-of-plane improper torsion potentials.

In (bio)molecular simulations, it is common to use periodic boundary conditions in order to simulate molecules in a realistic environment while allowing for the lower computational cost. When the box is big enough, the molecules will not interact with their periodic copies. The periodic boundary condition is closely related to the nonbonded interactions in which should be summed over all neighbours in the infinite periodic system. The system is normally coupled to a thermostat during integration to maintain room temperature. In addition, the total pressure in the system can be adjusted through scaling the simulation box size. The most troublesome part in the simulation is the computation of nonbonded interaction because every possible pair within the nonbonded cutoff radius has to be evaluated for each time step. It is essential to improve the simulation performance by extending the time step. In most of the simulations, the bond vibrations are not part of the interest and can be removed by introducing the bond constraint algorithm such as LINCS or SHAKE [96]. The algorithm constraints make it possible to extend time steps.

Another way to increase the time step, and thus extend the sampling, is the reduction of degrees of freedom, as done in coarse-grained (CG) simulations [97]. Atomistic simulations have limitation for application only small length and time scales. However, simulation for lipid solution, which has long chain structure and complicated system require one model which more computationally efficient. CG model simplified a small group of atoms into the effective interaction site [98, 99]. The reduction of the number of degree of freedom makes CG model can be applied to larger systems and longer time scales thus the significantly computational faster. CG model can be studied using a different of techniques for instance MD, Monte Carlo simulation and dissipative particle dynamics (DPD) [97, 100]. DPD methods allow lipid to be modelled as the soft beads and interact each other with a combination of repulsive, dissipative and random forces [97]. The CG model continues to be developed until now with some main aspects to achieve. These main aspects are speed, accuracy, applicability and versatility. The speed can be achieved by including only short-range interaction whereas the accuracy can be increased by mimic CG with atomistic simulation as much as possible. The applicability is reached by the use of force field and the versatility is referred as the force field can describe the detailed structure of the molecules [97].

2.3 Coarse-grain Martini Force Field Model

The Martini CG model uses basic rule four to one interaction site (bead). This means the four heavy atoms are represented by a single interaction centre [101-104].The four to one mapping was chosen because of its computational efficiency and chemical representability for molecules [102]. However, in ring compounds, a two or three to one mapping is used to preserve the planar geometry. The treatments or ring structures will be described more thoroughly later in the section. There are four main types of interaction sites with the symbol: polar (P), nonpolar (N), apolar (C) and charged (Q). Each of particle type has a number of subtypes yielding to the most specific representation of the chemical nature of atomic structure. The total number of the subtype is 18. The Martini CG model is considered as a good model because its transferability of the models made possible by the use if building block approach. The Martini CG model can be used for the different range of application. It does not need to be reparametrized for a different model. We can just match our building block of the structure with the building block structure of the model [101, 102].

The Lennard-Jones (LJ) potential energy is used for describing the non-bonded interaction [97, 101, 102]. The nonbonded of interaction sites *i* and *j* can be described as

$$U_{LJ}(r) = 4 \in_{ij} \left[\left(\frac{\sigma_{ij}}{r} \right)^{12} - \left(\frac{\sigma_{ij}}{r} \right)^{6} \right] \quad (39)$$

The σ_{ij} represent the closest distance of approach between two particles and $\in_{ij}$ describes the strength of their interaction. The interaction strength $\in$ of each interaction level can be divided into sub-levels as follows: O, $\in$ = 5.6 kJ/mol; I, $\in$ = 5 kJ/mol; II, $\in$ = 4.5 kJ/mol; III, $\in$ = 4 kJ/mol; IV, $\in$ = 3.5 kJ/mol; V, $\in$ = 3.1 kJ/mol; VI, $\in$ = 2.7 kJ/mol; VII, $\in$ = 2.3 kJ/mol; VIII, $\in$ = 2 kJ/mol; and I, $\in$ = 2.0 kJ/mol (with σ = 0.62 nm). The level of interaction is defined as strong polar interaction in the level I interaction. Whereas level IX can be described as the interaction between charged particles and a very apolar medium [101].

Moreover, there is another interaction for charged groups which called a shifted Coulombic potential energy function.

$$U_{el}(r) = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r} \quad (40)$$

with a relative dielectric constant ϵ_r= 15 for explicit screening.

The bonded interaction of the connected sites is described by a weak harmonic potential

$$V_{bond}(R) = \frac{1}{2} K_{bond}(R - R_{bond})^2 \quad (41)$$

The equilibrium distance $R_{bond} = \sigma = 4.7$ nm and a force constant of K_{bond} = 1250 kJ mol^{-1} nm^{-2} are standard parameters for alkyl chains. The LJ is excluded between bonded particles to avoid the too high forces between neighbouring bonded atom.

The ring particles are modelled slightly differently. For maintaining the geometry of small ring compounds, the basic rule four to one interaction site is incapable. There is a necessity to include as many CG sites in order to keep the ring geometry. Therefore, the mapping of two or three to one mapping of ring atoms into CG beads is the right way to do [101, 103]. The detailed mapping and appropriate geometry are better to mimic the geometry of small compounds for instance cyclohexane and benzene. There is a label “S” to denote the ring particles. The effective interaction size and strength for ring interaction is reduced comparing with the normal one. The σ in LJ potential is 0.43 nm instead of 0.47 nm and the ϵ is set to 75% of the original value. For instance the C_1 for the normal type has ϵ = 3.5 kJ mol $^{-1}$ and σ = 0.47 nm whereas the SC_1 type has a LJ potential with ϵ = 0.75 x 3.5 kJ mol^{-1} and σ = 0.43 nm.

3. Computational Details

3.1 Coarse-grain models for PCBM, DMPCBM and DMPCHp

1. PCBM

CG MD simulation based on the Martini CG force field were used for the models in the simulations. The CG model of PCBM uses the Martini C_{60} fullerene model developed by Monticelli [105] and available at the website [106]. The C_{60} fullerene model developed by Monticelli use the 16 beads representation of C_{60} fullerene. Comparison with C_{60} structure, PCBM has an addition of phenyl butyric acid methyl ester side chain part, as shown in Figure 13.

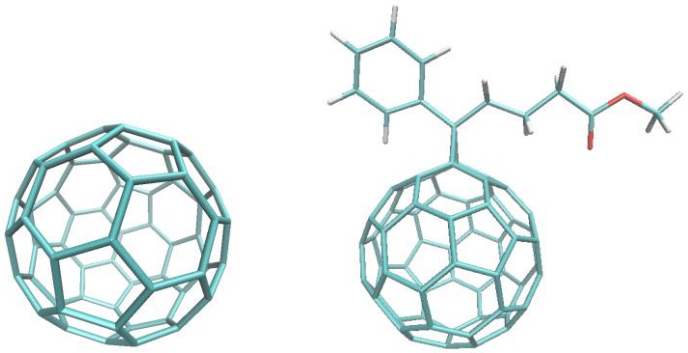
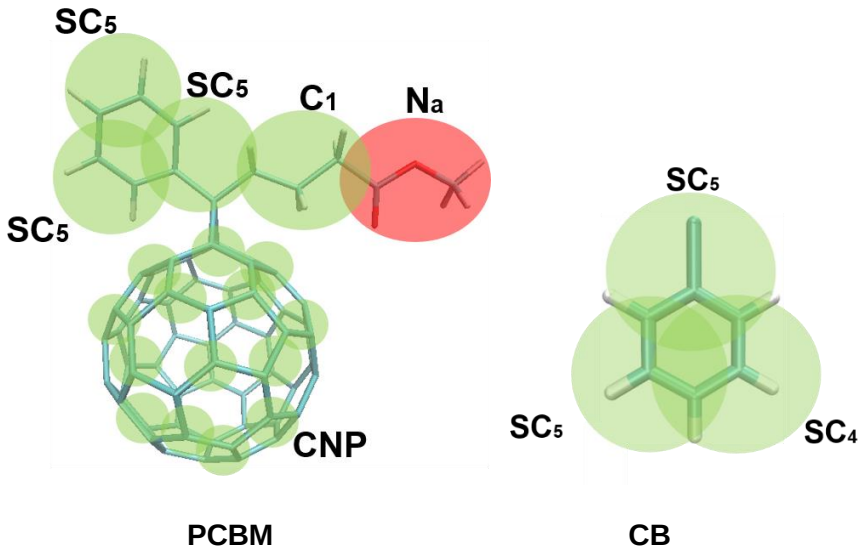
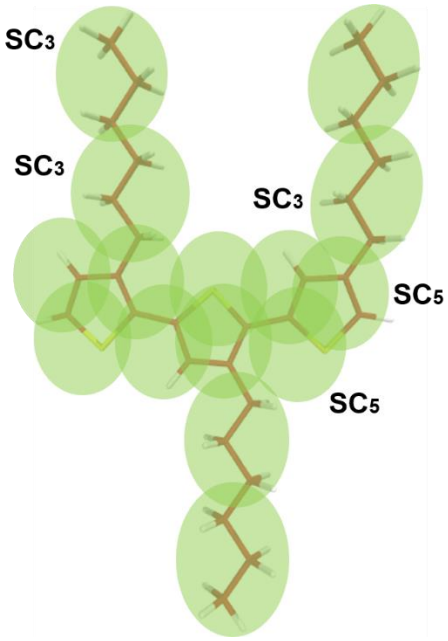


Fig. 13 Fullerene C_{60} (left) and PCBM structure (right).

The side chain of PCBM is represented by five interaction sites which consist of three Martini SC_5 beads referring to the phenyl structure whereas the butyric acid methyl ester side chain is represented with one C_1 (butane) and N_a (ester) beads. Figure 14 describe all the beads representation for PCBM, P3HT and CB.





P3HT

Fig. 14 PCBM, P3HT and chlorobenzene (CB) beads representation

P3HT polymer CG Martini model were used for our simulation. The beads types which define the nonbonded interaction were chosen based on the free energy of partitioning data. The three SC_5 beads represent the thiophene ring and two SC_3 beads represent the side chain. Chlorobenzene (CB) was used as the solvent for P3HT and PCBM because of its popularity among other solvents [107]. The Martini model has been created based on Martini benzene model. There are two SC_4 beads and one SC_5 bead describing the nonbonded interaction in CB molecules [107].

2. DMPCBM

In general, DMPCBM has similar structure except for the difference in phenyl structure of side chain. Modifying the phenyl side chain beads for DMPCBM was performed for making a CG model for DMPCBM. The phenyl side-chain structure of DMPCBM is 1,3-dimethoxybenzene (DMB). There was no need to rerun the whole fullerene derivative as all that change is only from benzene to DMB. Therefore, instead of performing the all-atom (AA) model simulation for all structure of DMPCBM, we performed atomistic simulation only for the side chain of the DMPCBM. Therefore, we perform the atomistic simulation for 1,3-dimethoxybenzene (DMB).

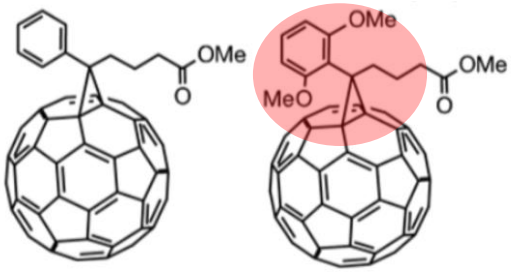


Fig. 15 PCBM structure (left) and DMPCBM structure (right)

Atomistic starting topology for DMB has been obtained from Automated Topology Builder (ATB) [108] based on GROMOS 53A6 force field [109]. The atomistic simulation was performed for one molecule DMB inside chlorobenzene (CB). The simulation was run for 50 ns. From the atomistic trajectory, we made the coarse-grained trajectory. Analysis of CG trajectory was performed and three bonds distance were obtained. These three bond distances will be put in the CG topology of DMPCBM and DMPCHp.

3. DMPCHp

DMPCHp has the similar structure as DMPCBM except in alkyl side chain structure. In DMPCBM, we have butyric acid methyl ester side chain whereas in DMPCHp we have heptane side chain (Figure 16. part highlighted in red). We have modified the alkyl side chain structure of DMPCBM by using the building block from reference [101]. The beads for pentane consisted of two beads of C_1 with bond length 0.47 nm and force constant 1250 kJ/mol.

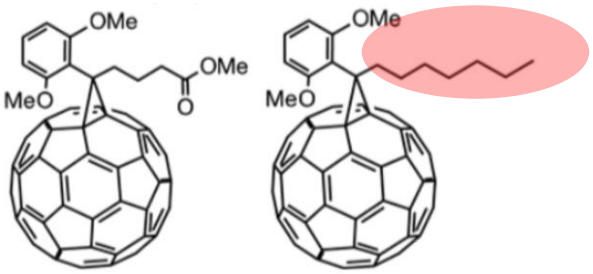


Fig. 16 DMPCBM structure (left) and DMPCHp structure (right).

3.2 Simulation parameter

The simulation parameters are based on the latest published parameters recommended for the Martini force field [110]. Simulation have been run using GROMACS v 5.0.4 [111]. The nonbonded interaction r_{cut} = 1.1 nm in combination with Verlet-neighborlist scheme. The temperature control is achieved by using velocity

rescale (V-rescale) [112] and thermostat using couling constant of 1 ps. The pressure is controlled with the Parrinello-Rahman barostat [113], with the pressure coupling of semiisotropic. The complete file for parameter can be seen in the Supporting Information (SI). All the simulation use the timestep between 20-30 fs. The energy minimization, the NVT and NPT equilibration were performed before run the production run.

3.3 Simulations in solution

A series of simulations of PCBM, DMPCBM and DMPCHp in CB were performed. The purpose of this simulation is to know the behaviour of PCBM, DMPCBM and DMPCHp and the probability of them to aggregate inside the solvent. Setting up simulation was started with making a box 8x8x8 nm^3. Five PCBM molecules were dissolved in CB solvent, as shown in Figure 17. Time length simulation was 300 ns. The same procedure was used for the simulation of DMPCBM and DMPCHp. Analysis of radial distribution function (RDF) was performed. In addition, we performed CG analysis from CG trajectory of the simulation. From sixteen beads representing PCBM, there will be two beads representing PCBM. The purpose of this RDF analysis was for finding the probability configuration arrangement of PCBM inside solvent. RDF analysis of fullerene C_{60} and side chain were performed. All of the simulation had been done using GROMACS 5.04 program.

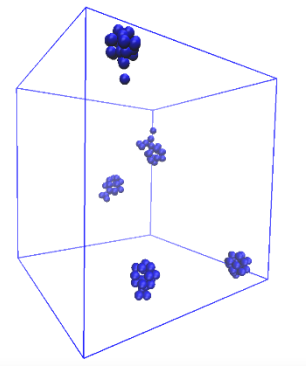


Fig. 17 Solvation box for simulation inside solvent

3.4 Solvent evaporation Simulation

Evaporation simulation was designed for mimicking as much as possible of experiments environment synthesis of bulk heterojunction films. Recently, Lee and Pao [80] developed an approach for solvent evaporation simulation. This is continued by Alessandri *et al.* [107] whom adopted a similar procedure and employed the Martini

force field. This approach is also taken in the present work. Solvent evaporation method was based on the spin coating process employed by experimentalist for making BHJ solar cells. General summary of the methods can be described as follows; in starting point, we will have a box which consisted of PCBM, P3HT and CB as the solvent. In every step of particular running simulation time, solvent molecules were removed until a dry thin film is obtained, as depicted in Figure 18 There are two variation size of the box were used. The boxes were 15x15x80 nm^3 and 20x20x80 nm^3. In addition, we have two different ratio of P3HT:fullerene; the first one is 0.7:1 and 1:0.8. The starting simulation box containing a mixture of P3HT (24mer): PCBM: CB (total concentration 70 mg/mL) with the corresponding ratio. From the starting solution, 1.25% of solvent molecules were randomly removed every time interval *t* until a dry thin film was obtained.

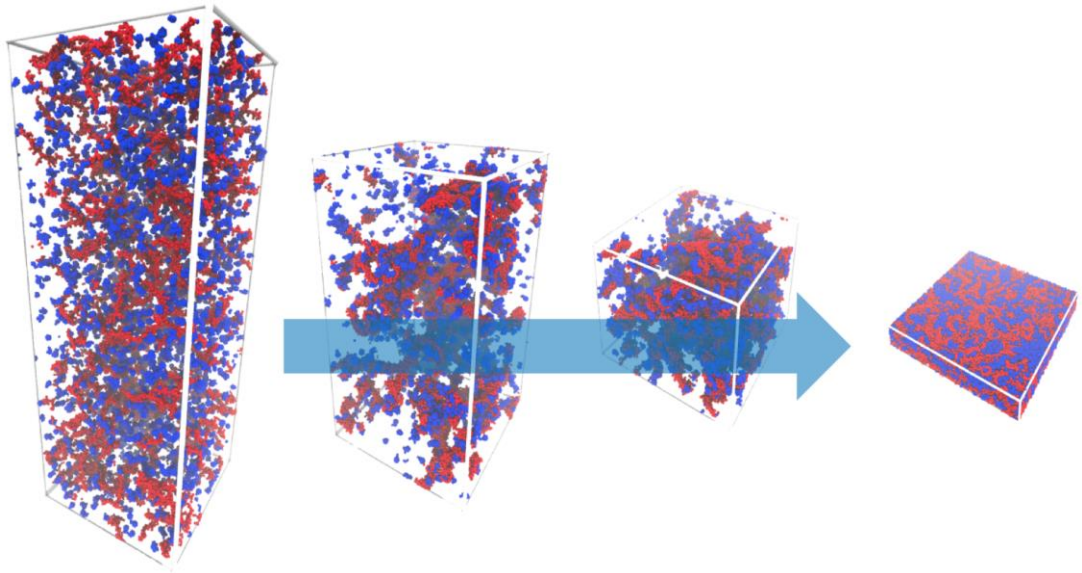


Fig. 18 Solvent evaporation simulation was adapted from Alessandri [107].

3.5 Morphological Analysis of fullerenes derivatives

Snapshot of final morphological appearances of the dry thin film was taken and analysis number of contacts were performed. The number of contacts of P3HT and PCBM can represent the interaction between polymer P3HT and PCBM.

The further analysis was computed using GROMACS tool, `gmx mindist`. This tool computes the distance between one group and a number of other groups. The minimum distance between any pair of atoms from the respective groups and the number of contacts for a given distance will be obtained. The systems have two group which are P3HT and PCBM. The `gmx mindist` calculate the number of contacts

between P3HT and PCBM within the distance of 0.6 nm. This distance was chosen because the radius of Martini beads is 0.27 nm. Therefore, any contacts between beads two group of P3HT and PCBM within the distance of 0.6 nm were computed.

4. Results and discussion

4.1 Coarse-grained (CG) model for PCBM, DMPCBM and DMPCHp

As mentioned above, the particular differences structure between PCBM and DMPCBM is in the phenyl structure of side chain. The type of beads has been modified for building new beads representation of DMPCBM. Figure 19 gives an explanation of the differences of beads representation in PCBM and DMPCBM. PCBM phenyl structure side chain is benzene which can be represented as three beads of SC_5. In comparison, DMPCBM phenyl structure part is 1,3- dimethoxybenzene which can be described as two beads of N_0 and one bead of SC_5. The modification type of beads is based on Marrink, *et al*. [101].

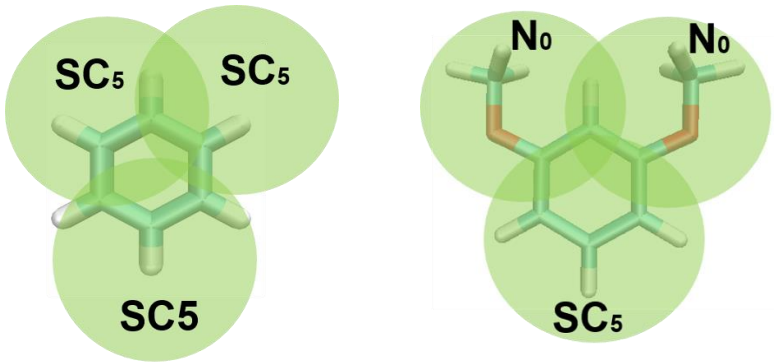


Fig. 19 Beads representation of PCBM (left) and DMPCBM (right)

The AA simulations were performed for the 1,3-dimethoxybenzene (DMB). The trajectory obtained from AA simulation was coarse-grained using the GROMACS tool, `gmx traj`. For making the CG from AA trajectory, we need the bead list file. The bead list file contains the information about the number of the beads from 0 until N-1 (for a molecule described in N beads). Analysis the CG trajectory were performed to extract bond distribution for three bonds in DMB molecule. GROMACS tool, `gmx distance` was used to calculate distances between the pairs of the centre of the mass for each bead. In DMB molecule, we have three beads, therefore we have three pairs of position. The 1st pair is N_0 and SC_5, the 2nd pair is SC_5 and SC5, lastly, the 3rd pair is N_0 and N_0. In the end of the analysis, we will get the three distances between the centre of mass of three DMB. The 1st and 2nd bond distance between the centre of mass of N_0

and SC_5 are 0.29 nm whereas the 3rd bond distance between the centre N_0 and N_0 is 0.36 nm. The bond distances are increased by 20% following the benzene model from AA to CG model. In AA model, the N_0 and SC_5 distance is 0.29 nm whereas in CG model is 0.36 nm. The bond distance of N_0 and N_0 in AA model is 0.36 nm whereas in CG is 0.45 nm.

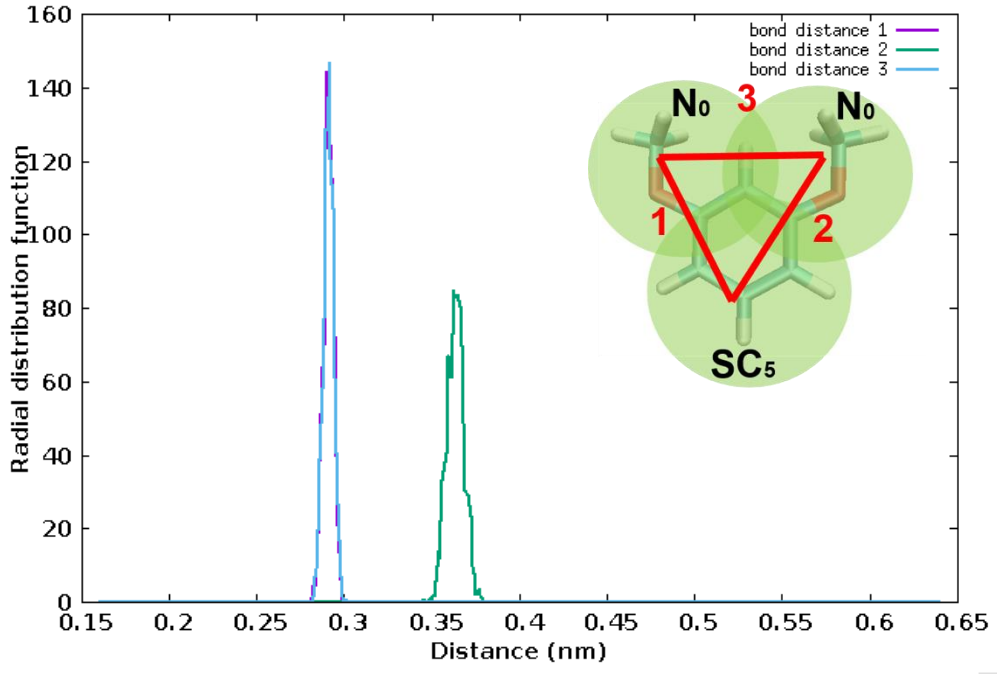


Fig. 20 Bond distance in the centre of mass of 1,3-dimethoxybenzene (DMB).

DMPCHp and DMPCBM have the similar structure in phenyl structure but different in alkyl side chain. Modification of butyric acid methyl ester side chain was conducted for building the DMPCHp model. DMPCBM has one C_1 and N_a bead whereas DMPCHp had two beads of C_1 as alkyl side chain. The C_1 bead is the bead for normal alkyl structure which has bond length 0.47 nm and force constant 1250 kJ mol^{-1} based on the paper of Marrink *et al*. [101].

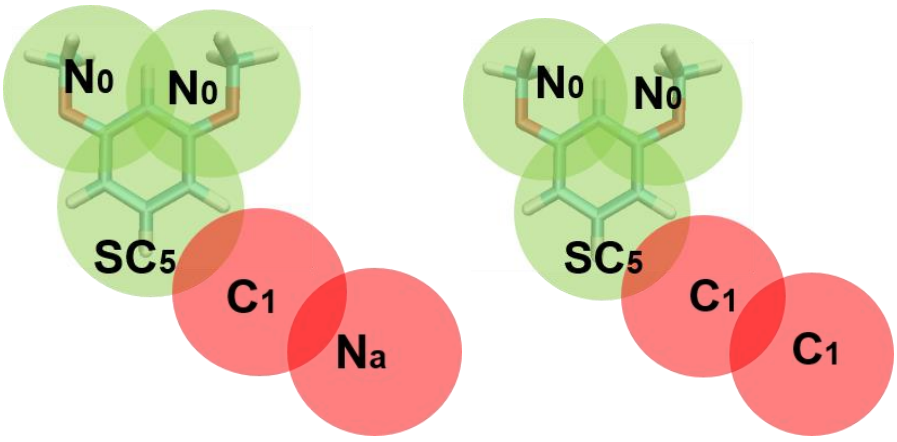


Fig. 21 Beads representation of DMPCBM (left) and DMPCHp (right).

4.2 Simulations in solution

Simulation of PCBM, DMPCBM and DMPCHp inside the solvent were conducted for studying the behaviour of three different fullerenes inside the solvent. This will also can give an indication whether three of fullerene will start to aggregate inside the solvent. The radial distribution function analysis was used for studying the behaviour of molecules inside the solvent. Further investigation was performed for analysis the radial distribution in which we performed CG analysis from CG trajectory of the simulation. The main idea of this investigation is making a CG analysis again from the CG trajectory. After running a simulation inside the solution, we will get the final trajectory of simulation, we refer this as the first trajectory. Then from this trajectory, we make a CG trajectory (refers as the second trajectory) which will be used for RDF analysis. In the first trajectory, PCBM has 21 beads, however, we did the remapping of beads so in the second trajectory we have PCBM with only two beads using GROMACS tool, `gmx traj`. One of the beads refers to the fullerene C60 of PCBM and the other one refers to the side chain part of PCBM. This analysis was done for all three fullerene derivatives.

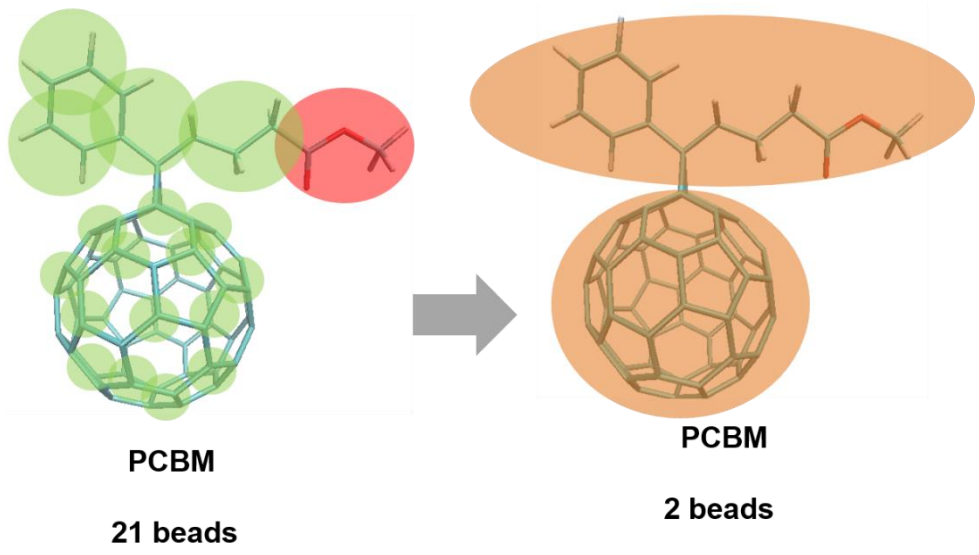


Fig. 22 The procedure of making CG trajectory from 21 beads into 2 beads.

The structure of liquids can be studied using radial distribution function (RDF). GROMACS tool, `gmx rdf` is used for calculating the RDF. One of the methods is to calculate the centre of mass (com) of the molecules to the closest molecules. The radial distribution function analysis was performed for two different aims. The first aim is performing RDF analysis to study the interaction between one fullerene C_{60} of PCBM

with the nearest another fullerene C_{60} of PCBM. For a distance less than 1 nm, there are no peaks because the two molecules cannot occupy the same space. The first peaks for all the three of different fullerene derivatives appear around 1.1 nm and the second peaks appear around 1.5 nm. Comparing the three different fullerene derivatives, DMPCBM has the lowest RDF among all (Figure 23). It means there is a weaker interaction between fullerene C_{60} of DMPCBM-DMPCBM comparing to fullerene C_{60} of PCBM-PCBM or DMPCHp-DMPCHp interaction. The weaker interaction between DMPCBM also gives the indication that DMPCBM will have higher solubility in CB because it is more polar than PCBM. Another particular behaviour from the three figures is PCBM and DMPCHp seem to have the similar characteristic with the RDF peaks. We expect this behaviour because if we compare both of structure particularly DMPCHp structure, the phenyl structure of DMPCHp become more polar however the alkyl chain structure becomes less polar therefore the two opposite effects probably cancel out.

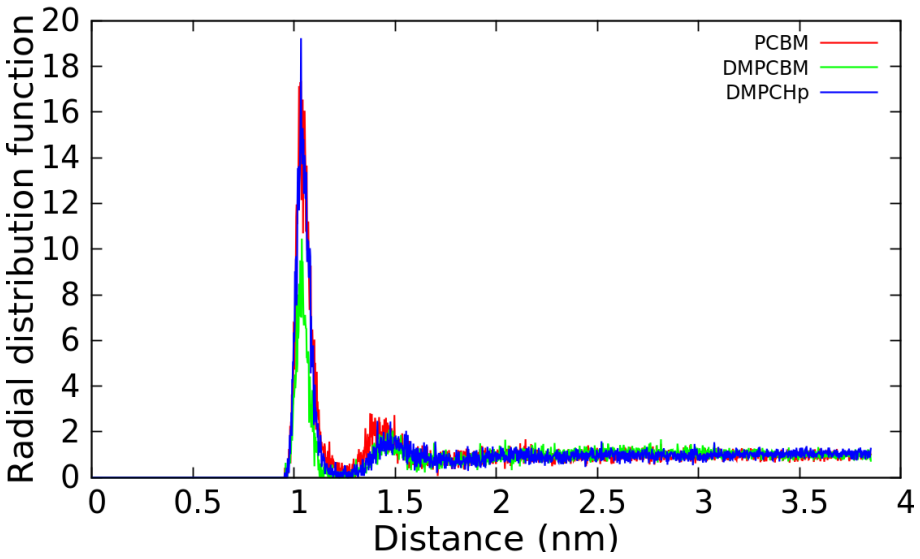


Fig. 23 Radial Distribution Function (RDF) of fullerene C_{60}

In addition, RDF analysis between fullerene C_{60} and side chain were studied. The main idea of this analysis is studying the probability of finding fullerene C_{60} of PCBM close to the side chain of the other nearest fullerene C_{60} of PCBM. PCBM has the highest peak among all the fullerene derivatives. Comparison between Figure 23 and Figure 24 shows that the RDF value of fullerene C_{60} is bigger than RDF value of fullerene C_{60} and side chain. It means the probability of finding fullerene C_{60} with the nearest another fullerene C_{60} close to each other is bigger than finding fullerene C_{60} with the nearest another side chain close to each other. We can conclude that the arranging order position of PCBM with the nearest PCBM will be fullerene C_{60} close to each other.

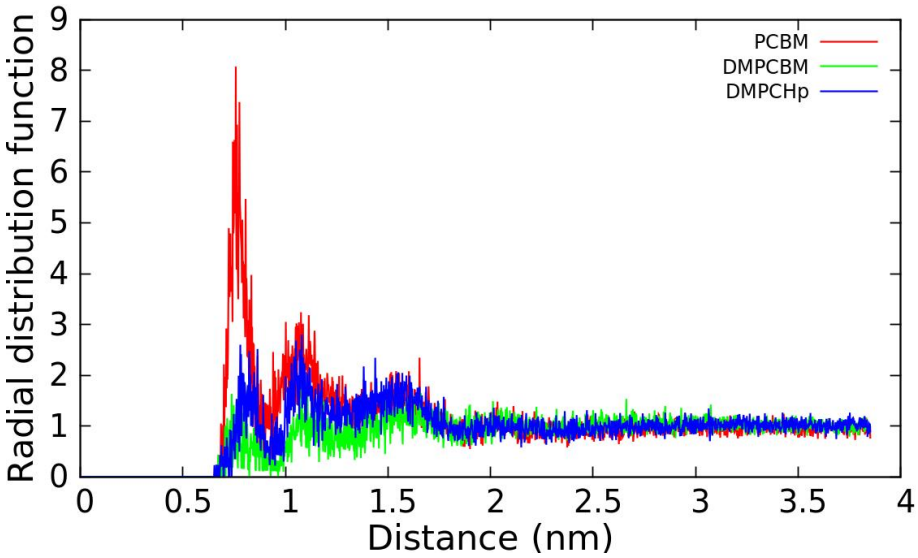


Fig. 24 Radial Distribution Function (RDF) of fullerene C_{60} and sidechain.

4.3 Solvent evaporation Simulation

Solvent evaporation simulation was performed for PCBM, DMPCBM and DMPCHp. We performed two different box variation for each fullerene. The boxes were 15x15x80 nm^3 and 20x20x80 nm^3. The starting box contains PCBM, P3HT inside CB solvent. In every 30 ns, we removed 1.25% CB solvent randomly. This can be noticed with the height of the box is getting smaller until the final thin film obtained. The final thickness of the film is about 5 nm and only PCBM (blue) and P3HT (red) left in the film. This solvent evaporation simulation was carried out to mimic experiment condition as much as possible. The morphological evolution of PCBM in solvent evaporation over time can be described in Figure 25 below.

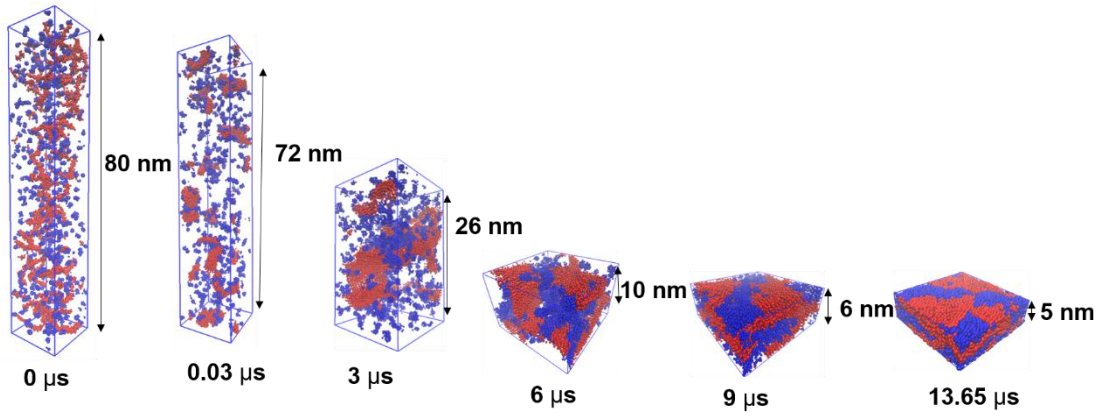


Fig. 25 Morphological evolution of PCBM in solvent evaporation simulation (red = P3HT, blue = fullerene)

4.4 Morphological Analysis of fullerene derivatives

We performed two different variations in the ratio of P3HT and PCBM. First ratio is 0.7:1 (P3HT:fullerene) and 1:0.8 (P3HT:fullerene). For each ratio, we performed the simulation for 15x15x80 nm^3 and 20x20x80 nm^3. The purpose of variation in the ratio is to know whether the ratio of P3HT and PCBM will affect the final morphology of the fullerenes derivatives. The morphological analysis of different fullerene derivatives was obtained from final morphology during solvent evaporation simulation. The final morphology of three simulations can give an explanation of phase separation happened during evaporation.

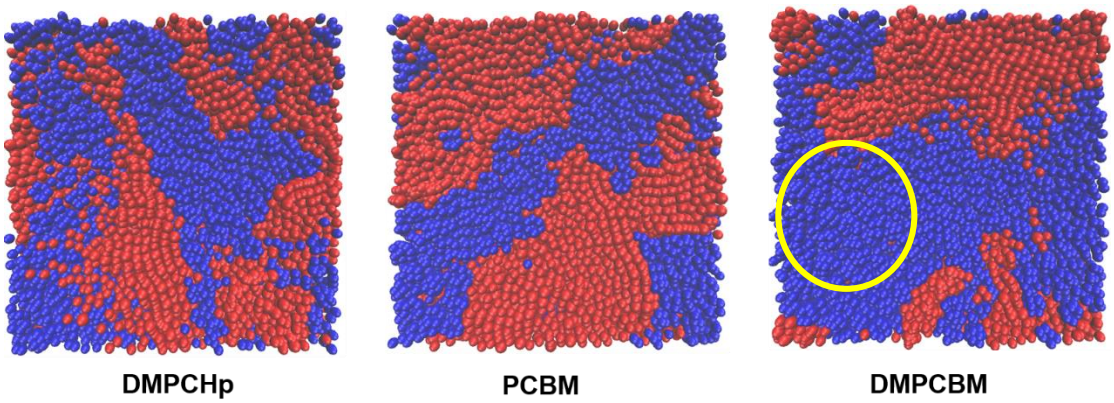
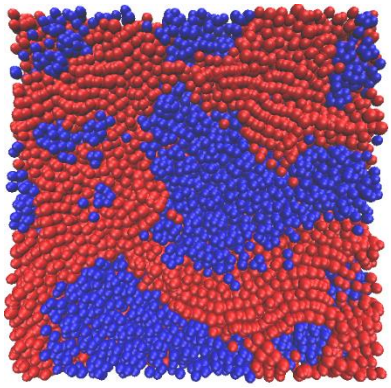
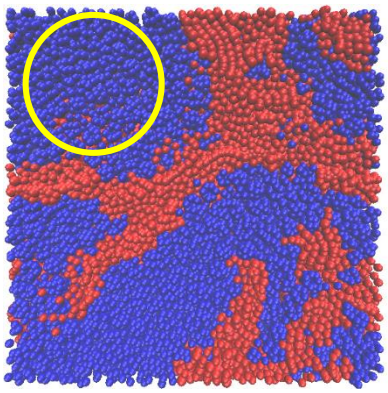


Fig. 26 Final morphological appearances of thin films DMPCHp, PCBM and DMPCBM using 0.7:1 (P3HT:fullerene) for 15x15x80 nm^3 box (red = P3HT, blue = fullerene). The area under yellow circle is the aggregation cluster of fullerene.

From three different sets of morphology in box 15x15x80 nm^3, DMPCBM (blue) tend to make a big cluster than the other two fullerenes indicated with a bigger blue cluster of P3HT. Whereas for PCBM and DMPCHp, the final of morphology cannot be distinguished which one has more aggregation. The simulation using 20x20x80 nm^3 boxes show similar results with 15x15x80 nm^3 boxes. The area inside yellow circle in Figure 26 and Figure 27 show the big area of DMPCBM. It means the DMPCBM make bigger aggregation than PCBM and DMPCHp. The number of contacts of P3HT-fullerene gives a quantitative measure about whether P3HT-fullerene is fully mixed with each other or phase separated. The bigger number of contacts P3HT-fullerene means less aggregation and less phase separation between P3HT-fullerene.



DMPCHp



DMPCBM

Fig. 27 Final morphological appearances of thin films DMPCHp, PCBM and DMPCBM using 0.7:1 (P3HT:fullerene) for 20x20x80 nm^3 box (red = P3HT, blue = fullerene). Area under yellow circle is the aggegration cluster of fullerene.

The number of contacts provides supporting information for morphological analysis of three different fullerene derivatives (see Table 17.). The number of contacts P3HT-DMPCBM is the least among the three of them. This means DMPCBM tends to aggregate more than PCBM and DMPCHp leading to more phase separation in the morphology of DMPCBM. Regarding DMPCHp, the number of contacts in P3HT-fullerene also decrease in comparison with PCBM. The difference in the number of contacts between P3HT-fullerene and P3HT-DMPCHp is less than the difference between P3HT-fullerene and P3HT-DMPCBM. This result is in agreement with Matsumoto *et al* [6] results in which those was found that DMPCBM also makes a cluster grain and do not mix as good with P3HT due to a lower affinity caused by the polarity mismatch. The number of contacts between 15x15x80 nm^3 and 20x20x80 nm^3 shows a similar trend in which DMPCBM has the least number of three fullerene derivatives.

Table 17. The number of contacts in three different fullerene derivatives in ratio 0.7:1

Number of contacts P3HT-fullerene	15x15x80	20x20x80
PCBM	20468	-
DMPCBM	16899	31074
DMPCHp	19058	36537

Furthermore, we had conducted the simulations to mimic the experimental ratio value which used for experimental research of Matsumoto *et al*. [81]. They use ratio 1:0.8 (P3HT:PCBM) which we adopt for our simulations. For each ratio, we performed the simulation for 15x15x80 nm^3 and 20x20x80 nm^3 boxes. We performed solvent evaporation analysis using ratio 1:0.8 (P3HT:fullerene) and instead of using 30 ns simulation time for each step, we used 15 ns simulation time for each step. Alessandri *et al*. [26] had conducted a series of simulation in which they concluded there is no

particular difference in simulation using 15 ns and 30 ns. The final morphological appearance of thin films for three different fullerene derivatives is presented in Figure 28 and Figure 29.

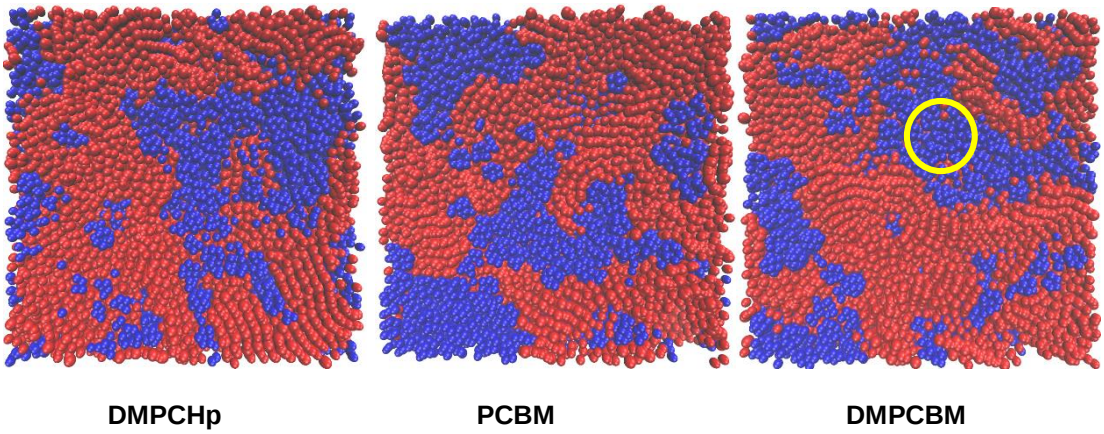


Fig. 28 Final morphological appearances of thin films DMPCHp, PCBM and DMPCBM using 1:0.8 (P3HT:fullerene) for 15x15x80 nm^3 box (red = P3HT, blue = fullerene). Area under yellow circle is the aggegration cluster of fullerene.

There are noticeable differences between the final morphological appearances of the thin film for the ratio 1:0.8 and 0.7:1. The two simulation has different ratio value in P3HT and PCBM. In the previous simulation, we had used ratio 0.7:1 (P3HT:fullerene), in which there are more PCBM molecules whereas in ratio 1:0.8 (P3HT:fullerene), there are more P3HT molecules. We have to keep in mind that in ratio 1:0.8, there are fewer fullerene molecules, therefore, the yellow circle in Figure 28 is clearly smaller than Figure 26 and Figure 27. Simulation with 20x20x80 nm^3 boxes (Figure 29) shows similar results where the area under yellow circle are smaller.

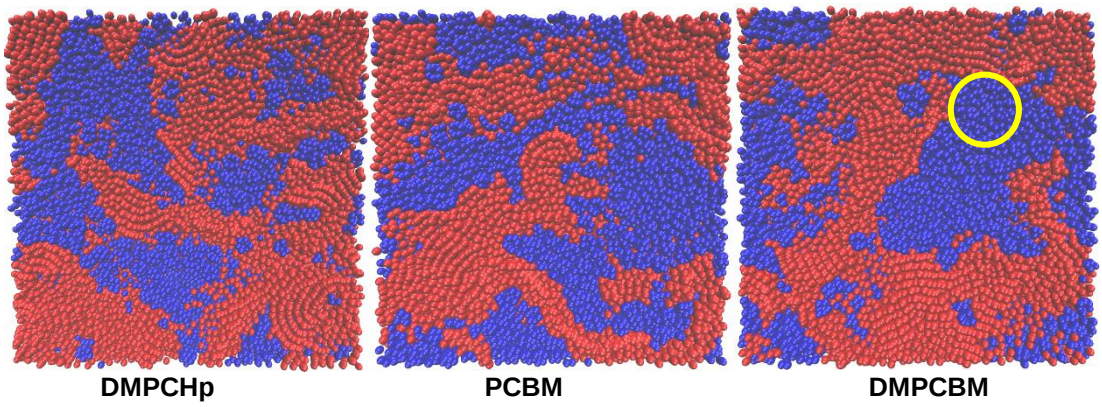


Fig. 29 Final morphological appearances of thin films DMPCHp, PCBM and DMPCBM using 1:0.8 (P3HT:fullerene) for 20x20x80 nm^3 box (red = P3HT, blue = fullerene). The area under yellow circle is the aggregation cluster of fullerene.

The aggregation behaviour can be explained using the number of contacts. In line with the previous ratio, DMPCBM with ratio 1:0.8 has the least number of contacts between the three fullerene derivatives. The number of contacts in DMPCBM with ratio

0.7:1 (P3HT:fullerene) has 31812 and DMPCBM with ratio 1:0.8 (P3HT:fullerene) has 31074. Both of the value are similar but have the different meaning. In the simulation with ratio 1:0.8, we have a lower amount of fullerene derivative than the previous simulation. Even though the simulation with ratio 1:0.8 have less amount of PCBM, it still manages to mix with the polymer P3HT. The number of contacts does not change massively so we can conclude that P3HT and PCBM are still managed to make the same number of contacts.

Table 18. The number of contacts P3HT-fullerene in three different fullerene derivatives in ratio 1:0.8.

Number of contacts P3HT-fullerene	15x15x80	20x20x80	20x20x80
PCBM	21189	42162	39807
DMPCBM	20305	31812	31213
DMPCHp	20487	36910	35785

We have calculated the standard deviation and standard error for the number of contacts using both boxes of 15x15x80 and 20x20x80. Since box 20x20x80 has a larger number of molecules, we have to renormalize the 15x15x80 to a 20x20x80. As a first approximation, we can multiply the number of contacts PCBM-P3HT with the ratio of volume between two boxes. Box 15x15x80 has volume 18000 nm^3 whereas box 20x20x80 has volume 32000 nm^3. The ratio of the volume is 1.78 (this value from 32000/18000). We get the new value for the number of contacts for 15x15x80 which now we can use to calculate the mean of the number of contacts. Furthermore, we can calculate the standard deviation (SD) and standard error (SE).

Table 19. Standard deviation and standard error for the number of contacts for ratio 0.7:1.

Number of contacts P3HT-fullerene	15x15x80 become 20x20x80	20x20x80	Mean (M)	Standard deviation (SD)	Standard error (SE)
PCBM	36388	-	-	-	-
DMPCBM	30043	31074	30558	516	365
DMPCHp	33881	36537	35209	1328	939

We have not get the number of contacts for the simulation of 20x20x80 box ratio 0.7:1 because the simulation is still running at this moment. We can conclude the number of contacts for PCBM is 36388, the number of contacts for DMPCBM is 30558 ± 365 and the number of contacts for DMPCHp is 35209 ± 939. The SD and SE values for both of data are small. Because of the limitation time, we have tried only for two batches of data however if we have more batch of data, the SD and SE will decrease. It

will be better if we have the same set of the box for calculating the standard deviation. However, in this case, we use a different kind of boxes and try to calculate the standard deviation which turns out is not the best solution.

Table 20. Standard deviation and standard error for the number of contacts for ratio 1:0.8.

Number of contacts P3HT-fullerene	15x15x80 become 20x20x80	20x20x80	20x20x80	Mean (M)	Standard deviation (SD)	Standard error (SE)
PCBM	37669	42162	39807	39879	2247	1297
DMPCBM	36098	31812	31213	33041	2664	1538
DMPCHp	36421	36910	35785	36372	564	326

Regarding this, we added one more batch calculation using a box of 20x20x80 for ratio 1:0.8 therefore, we have the three batch of data. We have used the same method for calculating SD and SE for ratio 1:0.8. We can conclude the number of contacts for PCBM is 39879 ± 1297, the number of contacts for DMPCBM is 33041 ± 1538 and the number of contacts for DMPCHp is 36372 ± 326. In line with the previous statement, we have tried only for three batches of data however if we the more batch of data included, the SD and SE will surely decrease.

4. 5 Analysis number of contacts over time

Analysis number of contacts over time were performed for studying the evolution of the number of contacts P3HT-fullerene in the solvent evaporation simulation step. In the beginning of solvent evaporation simulation, it is clearly less number of contacts P3HT-fullerene. However, with the evaporation of the solvent, there is an increasing number of contacts P3HT-fullerene. We have calculated the percentage of a number of the contacts P3HT-fullerene over the sum of all number of contacts. When we calculated the number of contacts, we did not only calculate the number of contacts of P3HT-fullerene but also calculated the number of contacts of P3HT-P3HT and fullerene-fullerene. The number of contacts of P3HT-P3HT is the number of contacts for polymer P3HT and P3HT whereas the number of the contacts of fullerene-fullerene is the number of the contact for between fullerene derivatives for instance between PCBM and other PCBM in one system. The sum of all the number of contacts is the sum of the number of contacts of P3HT-fullerene, P3HT-P3HT and PCBM-PCBM.

Table 21. The number of contacts for ratio 0.7:1 of the 15x15x80 box.

	PCBM	DMPCBM	DMPCHp
The number of contacts of P3HT-fullerene	20468	16899	19058
The number of contacts of P3HT-P3HT	134730	138020	135199
The number of contacts of fullerene-fullerene	80580	77176	76392
Sum of the number of contacts	235778	232095	230649
Percentage (%) P3HT-fullerene	8.68	7.28	8.26

First, we discuss the number of contacts over time for ratio 0.7:1 for the 15x15x80 box. The percentage (%) P3HT-fullerene is the number of contacts of P3HT-fullerene divided by the sum of the number of contacts. Then, we make a graphic from the percentage (%) versus simulation time. Figure 30 gives the representation of the evolution of the P3HT-fullerene number of contacts over the simulation time. In the beginning, the solvents are still in the system, therefore the number of contacts of P3HT-fullerene is rather small. However as the simulation is running, we remove the solvent CB and the number of contacts of P3HT-fullerene increase steadily. The simulation stop when all of the solvents is removed from the system and the thin film obtained.

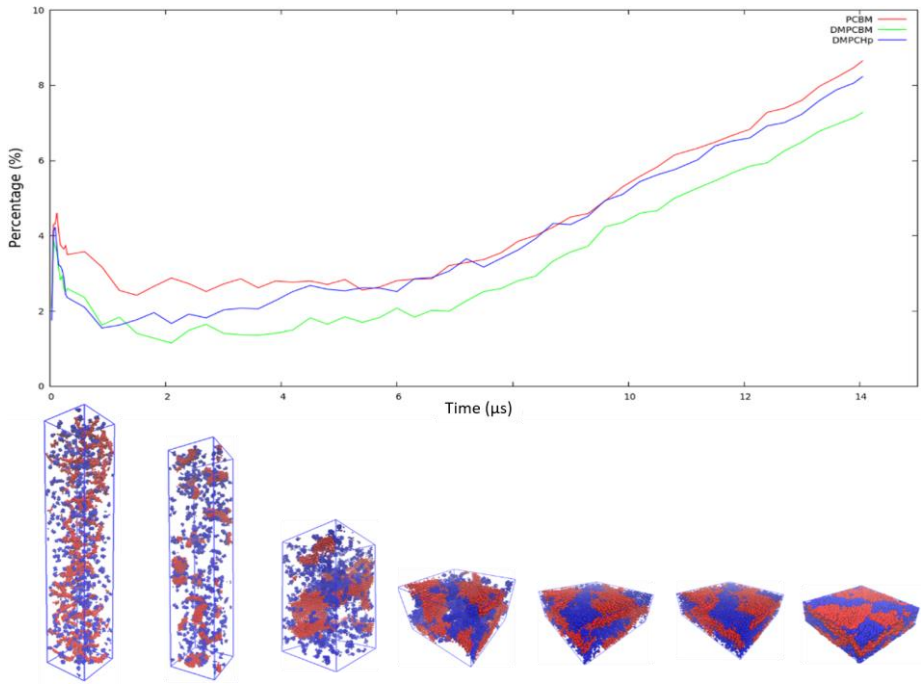


Fig. 30 The percentage (%) P3HT-fullerene number of contacts for ratio 0.7:1 of 15x15x80 box.

The comparison between three different fullerenes shows that DMPCBM has the lowest percentage (%) P3HT-fullerene among them which is 7.28% (Figure 30, green line). The difference of the percentage (%) between DMPCBM and PCBM (1.4%) or DMPCHp and DMPCBM (0.98%) are distinguishable with the difference more than 1%.

However, the difference between the percentage (%) between DMPCBM and PCBM (0.42%) is less than 1% in which the difference is rather small. The explanation of this number is similar to the explanation of Figure 13 when we talked about RDF results. The comparison of PCBM and DMPCHp structure show that even though the phenyl structure of DMPCHp become more polar but the alkyl chain structure become less polar therefore the opposite effects presumably compensate each other.

Table 22. The number of contacts for ratio 1:0.8 of 15x15x80 box.

	PCBM	DMPCBM	DMPCHp
The number of contacts of P3HT-fullerene	21189	20305	186547
The number of contacts of P3HT-P3HT	186547	187735	187236
The number of contacts of fullerene-fullerene	57486	54665	54096
Sum of the number of contacts	265222	262705	261819
Percentage (%) P3HT-fullerene	7.98	7.72	7.82

The same method was applied to calculate the sum of the number of contacts and the percentage (%) P3HT-fullerene for ratio 1:0.8 of the 15x15x80 box. DMPCBM has the least percentage (%) number of contacts of P3HT-fullerene than PCBM and DMPCHp. However, the difference between the percentage (%) of DMPCBM and PCBM (0.26 %) or DMPCHp and PCBM (0.1%) are rather small and both less than 1%. It is shown from the three line of fullerene derivatives (Figure 31) do not separate clearly. Because of these results, we should check and compare the results for simulation using 15x15x80 and 20x20x80 box.

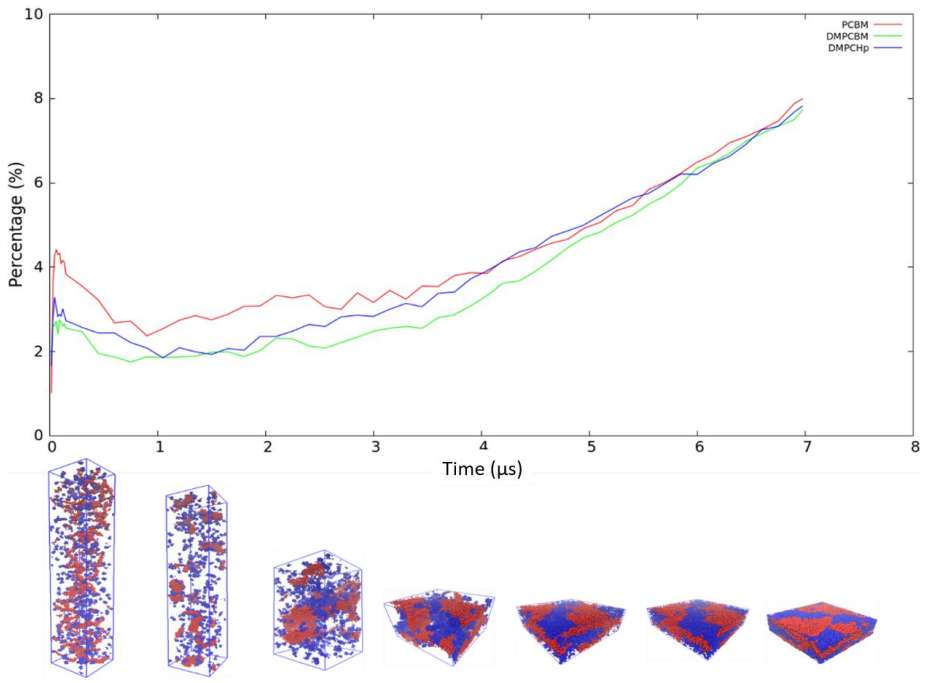


Fig. 31 The percentage (%) P3HT-fullerene number of contacts for ratio 1:0.8 of 15x15x80 box.

The summary of the percentage (%) P3HT-fullerene number of contacts ratio 1:0.8 of 20x20x80 is presented in Table 23. Analogous with results of 15x15x80 box, the results of 20x20x80 show that DMPCBM has the lowest percentage (6.81%) P3HT-fullerene among three derivatives. However, the difference between the percentage of DMPCBM and PCBM (2.14%) or DMPCHp and PCBM (1%) are bigger and both more than 1% in comparison with results of 15x15x80 box. The three lines in Figure 32 are clearly separated in the final step of simulation in contrast with Figure 31.

Table 23. The number of contacts for ratio 1:0.8 of 20x20x80 box.

	PCBM	**DMPCBM**	**DMPCHp**
The number of contacts of P3HT-fullerene	42162	31812	36910
The number of contacts of P3HT-P3HT	328542	336637	331154
The number of contacts of fullerene-fullerene	100220	98587	95994
Sum of the number of contacts	470924	467036	464058
Percentage (%) P3HT-fullerene	8.95	6.81	7.95

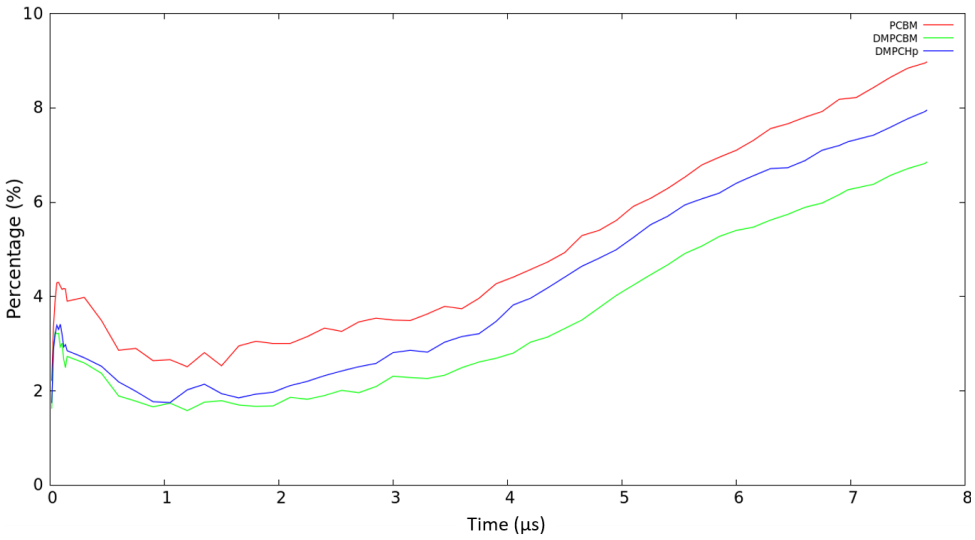


Fig. 32 The percentage (%) P3HT-fullerene number of contacts for ratio 1:0.8 of 20x20x80 box.

We had added another one batch of simulation of 20x20x80. Complementary with previous results, the second batch of 20x20x80 gives the similar results for the percentage of P3HT-fullerene. DMPCBM has the lowest number of contacts of P3HT-fullerene which is 6.71%, similar to previous results. The difference between the percentage of DMPCBM and PCBM (1.77%) or DMPCHp and PCBM (0.75%) are bigger than the results of 15x15x80 box.

Table 24. The number of contacts for ratio 1:0.8 for the 20x20x80 box (second batch)

	PCBM	DMPCBM	DMPCHp
The number of contacts of P3HT-fullerene	39807	31213	35785
The number of contacts of P3HT-P3HT	327832	335081	329819
The number of contacts of fullerene-fullerene	101774	98976	96833
Sum of the number of contacts	469413	465270	462437
Percentage (%) P3HT-fullerene	8.48	6.71	7.73

Both of the results from the simulation of 20x20x80 boxes have a distinguishable line. It is showed with the three of lines separated in Figure 32 and Figure 33. The results from 20x20x80 are more reliable than the results from 15x15x80 because of more consistency in term percentage of the P3HT-fullerene number of contacts. Therefore, to explain the evolution of the number of contacts of P3HT-fullerene for ratio 1:0.8, it is better using the 20x20x80 box. Nevertheless, the calculation for ratio 1:0.8 of 15x15x80 had done only one time so we cannot exclude the results for calculating the standard deviation (SD) and standard error (SE) of the number of contacts.

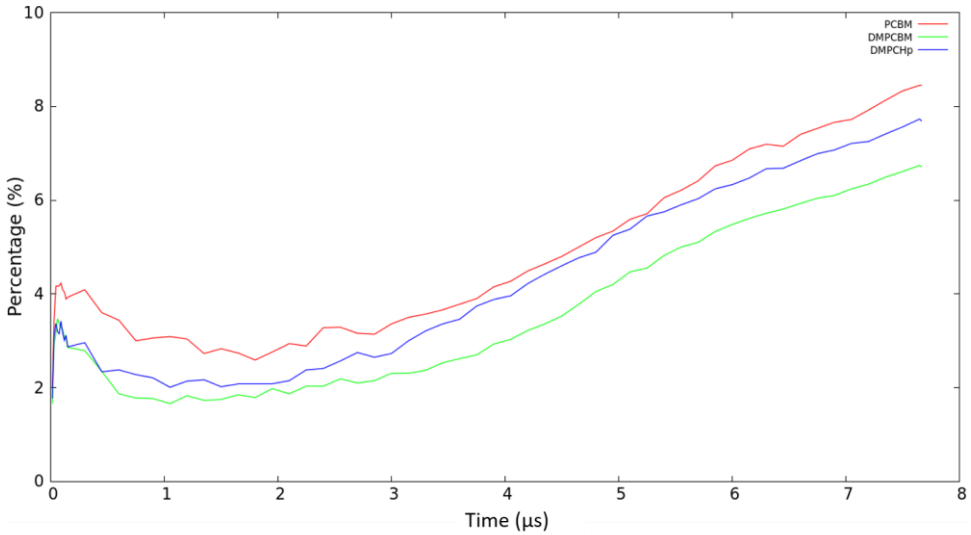


Fig. 33 The percentage (%) P3HT-fullerene number of contacts for ratio 1:0.8 of 20x20x80 box (second batch).

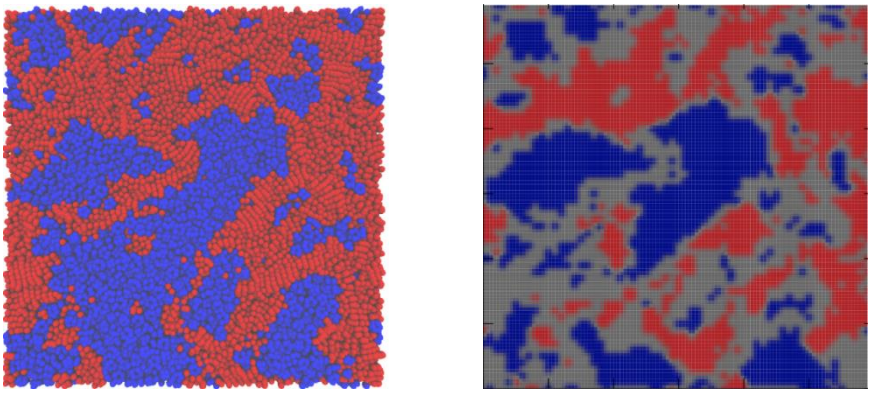
5. Conclusion

Our current works show that we can modify the Martini CG model for PCBM to be used for other fullerenes derivatives such as DMPCBM and DMPCHp. From the evaporation simulation, it has been shown DMPCBM tend to aggregate more than PCBM and DMPCHp. This is probably due to the lower affinity of DMPCBM and

polymer P3HT. The number of contacts P3HT-fullerene decrease with the increasing of the polarity of fullerene derivatives. DMPCBM have less number of contacts in P3HT-fullerene among the three of them. These results are in agreement with the experimental findings of Matsumoto et *al.* [81]. There are noticeable differences between the final morphological appearances of the thin film for the ratio 1:0.8 and 0.7:1. However, the number of contacts for both of the ratio does not change massively. We can conclude that even though we have a lower amount of fullerene derivative, it still manages to mix with the polymer P3HT.

6. Further Work

From our current work, we only performed quantitative morphological analysis of PCBM, DMPCBM and DMPCHp by using the number of contacts. In the future, we want to perform more quantitative morphological analysis of PCBM, DMPCBM and DMPCHp using spatial discretization scheme methods (Figure 34). By using this method, we hope we can calculate the size of the clusters for P3HT and PCBM.



(a) **(b)**

Fig. 34 Spatial discretization scheme (a) original morphology (b) after remapping.

In addition, we want to do several more batch of calculation to reduce the SD and SE of the number of contacts P3HT-fullerene to obtain a definitive picture of the impact of fullerene polarity on the P3HT:fullerene mixing.

References

1. Harding, M.M., M.W. Nowicki, and M.D. Walkinshaw, *Metals in protein structures: a review of their principal features.* Crystallography Reviews, 2010. **16**(4): p. 247-302.
2. Siegbahn, P.E.M. and M.R.A. Blomberg, *Transition-Metal Systems in Biochemistry Studied by High-Accuracy Quantum Chemical Methods.* Chem. Rev, 2000. **1000**: p. 421-437.
3. Tamames, J.A. and M.J. Ramos, *Metals in proteins: cluster analysis studies.* J Mol Model, 2011. **17**(3): p. 429-42.
4. Broderick, J.B., *Coenzymes and Cofactors.* Encyclopedia of Life Science, 2001.
5. Baker, H.M., B.F. Anderson, and E.N. Baker, *Dealing with iron: common structural principles in proteins that transport iron and heme.* Proc Natl Acad Sci U S A, 2003. **100**(7): p. 3579-83.
6. Liu, J., S. Chakraborty, P. Hosseinzadeh, Y. Yu, S. Tian, I. Petrik, A. Bhagi, and Y. Lu, *Metalloproteins containing cytochrome, iron-sulfur, or copper redox centers.* Chem Rev, 2014. **114**(8): p. 4366-469.
7. Ho, J., M.L. Coote, C.J. Cramer, and D.G. Truhlar, *Theoretical Calculation of Reduction Potentials*, in *Organic Electrochemistry: Revised and Expanded*, O. Hammerich and B. Speiser, Editors. 2012, CRC Press, Taylor and Francis Group. p. 229-259.
8. Arumugam, K. and U. Becker, *Computational Redox Potential Predictions: Applications to Inorganic and Organic Aqueous Complexes, and Complexes Adsorbed to Mineral Surfaces.* Minerals, 2014. **4**(2): p. 345-387.
9. Su, N.Q., I.Y. Zhang, J. Wu, and X. Xu, *Calculations of ionization energies and electron affinities for atoms and molecules: A comparative study with different methods.* Frontiers of Chemistry in China, 2012. **6**(4): p. 269-279.
10. Kulik, J. and N. Marzari, *Electronic Structure and Reactivity of Transition Metal Complexes*, in *Fuel Cell Science: Theory, Fundamentals, and Biocatalysis*, W. Wieckowski and J.K. Nørskov, Editors. 2010, John Wiley & Sons, Inc.
11. Riley, K.E. and K.M. Merz, Jr., *Assessment of Density Functional Theory Methods for the Computation of Heats of Formation and Ionization Potentials of Systems Containing Third Row Transition Metals.* J. Phys. Chem. A, 2007. **111**(27): p. 6044-6053.

12. Riley, K.E., B.T. Op't Holt, and K.M. Merz, Jr., *Critical Assessment of the Performance of Density Functional Methods for Several Atomic and Molecular Properties.* J Chem Theory Comput, 2007. **3**(2): p. 407-433.
13. Uudsemaa, M. and T. Tamm, *Density-Functional Theory Calculations of Aqueous Redox Potentials of Fourth-Period Transition Metals.* J. Phys. Chem. A, 2003. **107**(46): p. 9997-10003.
14. Li, S., J.S. Hennigan, D.A. Dixon, and K.A. Peterson, *Accurate Thermochemistry for Transition Metal Oxide Clusters.* J. Phys. Chem. A, 2009. **113**(27).
15. Zhou, Z., M.L. Steigerwald, R.A. Friesner, L. Brus, and M.S. Hybertsen, *Structural and chemical trends in doped silicon nanocrystals: First-principles calculations.* Physical Review B, 2005. **71**(24).
16. Jensen, F., *Introduction to Computational Chemistry* Vol. 2. 2007, England: John Wiley & Sons Ltd.
17. Cramer, C.J., *Essentials of Computational Chemistry Theories and Models* John Wiley & Sons Ltd, 2004.
18. Raghavachari, K., G.W. Trucks, J.A. Pople, and M. Head-Gordon, *A Fifth Order Pertubation Comparison of Electron Correlation Theories.* Chem. Phys. Lett., 1989. **157**(479).
19. Truhlar, D.G., *Basis-set Extrapolation.* Chemical Physics Letters, 1998. **294**: p. 45-48.
20. Halkier, A., T. Helgaker, P. Jørgensen, W. Klopper , H. Koch, J. Olsen, and A.K. Wilson, *Basis-set convergence in correlated calculations on Ne, N , and H2O.* Chem Phys Letters, 1998. **286**: p. 243-252.
21. Helgaker, T. and W. Klopper *Basis-set convergence of correlated calculations on water.* J. Chem. Phys, 1997. **106**(23).
22. Burke, K., *Perspective on density functional theory.* J Chem Phys, 2012. **136**(15): p. 150901.
23. Kohn, W. and L.J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects.* Physical Review, 1965. **140**(4A): p. A1133-A1138.
24. Tsipis, A.C., *DFT flavor of coordination chemistry.* Coordination Chemistry Reviews, 2014. **272**: p. 1-29.
25. Perdew, J.P., *Jacob's ladder of density functional approximations for the exchange-correlation energy.* 2001. **577**: p. 1-20.
26. Perdew, J.P., *Climbing the ladder of density functional approximations.* MRS Bulletin, 2013. **38**(09): p. 743-750.

27. Perdew, J.P., A. Ruzsinszky, J. Tao, V.N. Staroverov, G.E. Scuseria, and G.I. Csonka, *Prescription for the design and selection of density functional approximations: more constraint satisfaction with fewer fits.* J Chem Phys, 2005. **123**(6): p. 62201.
28. Cohen, A.J., P. Mori-Sanchez, and W. Yang, *Challenges for density functional theory.* Chem Rev, 2012. **112**(1): p. 289-320.
29. Becke, A.D., *Density-functional thermochemistry. III. The role of exact exchange.* The Journal of Chemical Physics, 1993. **98**(7): p. 5648.
30. Zhao, Y. and D.G. Truhlar, *The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals.* Theoretical Chemistry Accounts, 2007. **120**(1-3): p. 215-241.
31. Frisch, M.J., G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M.J. Bearpark, J. Heyd, E.N. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A.P. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N.J. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, and D.J. Fox, *Gaussian 09*. 2009, Gaussian, Inc.: Wallingford, CT, USA.
32. Brás, N.F., M.A.S. Perez, P.A. Fernandes, P.J. Silva, and M.J. Ramos, *Accuracy of Density Functionals in the Prediction of Electronic Proton Affinities of Amino Acid Side Chains.* Journal of Chemical Theory and Computation, 2011. **7**(12): p. 3898-3908.
33. Jurecka, P. and P. Hobza, *On the convergence of the term ($\langle\langle\Delta E\rangle\rangle$_((CCSD(T)-MP2))^) for complexes with multiple H-bonds.* Chemical Physics Letters 2002. **365**: p. 89-94.

34. Vosko, S.H., L. Wilk, and M. Nusair, *Accurate Spin Dependent electron liquid correlation energies for local spin density calculation: A Critical Analysis.* Can. J. Phys, 1980. **58**: p. 1200.
35. Becke, A.D., *Density-functional thermochemistry. IV. A new dynamical correlation functional and implications for exact-exchange mixing.* The Journal of Chemical Physics, 1996. **104**(3): p. 1040.
36. Karton, A., A. Tarnopolsky, J.F. Lamere, G.C. Schatz, and J.M.L. Martin, *Highly Accurate First-Principles Benchmark Data Sets for the Parametrization and Validation of Density Functional and Other Approximate Methods. Derivation of a Robust, Generally Applicable, Double-Hybrid Functional for Thermochemistry and Thermochemical Kinetics.* J. Phys. Chem. A, 2008. **112**: p. 12868-12886.
37. Perdew, J.P., P. Ziesche, and H. Eschrig, *Electronic structure of solids' 91*. Vol. 11. 1991: Akademie Verlag, Berlin.
38. Zhao, Y., B.J. Lynch, and D.G. Truhlar, *Development and Assessment of a New Hybrid Density Functional Model for Thermochemical Kinetics.* J. Phys. Chem. A, 2004. **108**(14).
39. Grimme, S., *Semiempirical hybrid density functional with perturbative second-order correlation.* J Chem Phys, 2006. **124**(3): p. 034108.
40. Boese, A.D. and J.M. Martin, *Development of density functionals for thermochemical kinetics.* J Chem Phys, 2004. **121**(8): p. 3405-16.
41. Kozuch, S., D. Gruzman, and J.M.L. Martin, *DSD-BLYP: A General Purpose Double Hybrid Density Functional Including Spin Component Scaling and Dispersion Correction.* 114, 2010. **48**: p. 20801-20808.
42. Zhao, Y., N.E. Schultz, and D.G. Truhlar, *Exchange-correlation functional with broad accuracy for metallic and nonmetallic compounds, kinetics, and noncovalent interactions.* J Chem Phys, 2005. **123**(16): p. 161103.
43. Schwabe, T. and S. Grimme, *Towards chemical accuracy for the thermodynamics of large molecules: new hybrid density functionals including non-local correlation effects.* Phys Chem Chem Phys, 2006. **8**(38): p. 4398-401.
44. Becke, A.D., *Density-Functional Exchange-Energy Approximation with Correct Asymptotic-Behavior.* Phys. ReV. A: At., Mol., Opt. Phys. , 1988. **38**: p. 3098–3100.
45. Perdew, J.P., *Density-Functional Approximation for the Correlation-Energy of the Inhomogeneous Electron-Gas.* Phys. ReV. B: Condens. Matter Mater. Phys. , 1986. **33** p. 8822–8824.

46. Perdew, J.P., *Erratum: Density-functional approximation for the correlation energy of the inhomogeneous electron gas.* Phys. Rev. B, 1986. **34**(10): p. 7406-7406.
47. Zhao, Y., N.E. Schultz, and D.G. Truhlar, *Design of Density Functionals by Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions.* J. Chem. Theory Comput.,, 2006. **2**(2).
48. Perdew, J.P., K. Burke, and M. Ernzerhof, *Generalized Gradient Approximation Made Simple.* Phys. Rev. Lett., 1996. **77**(18): p. 3865-3868.
49. Perdew, J.P., *Unified Theory of Exchange and Correlation Beyond the Local Density Aproximation. In Electronic DFT Functionals Benchmarking* J. Chem. Theory Comput., 1991. **Vol. 6, No. 8**: p. 2010 2291.
50. Lee, C., W. Yang, and R.G. Parr, *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density.* Phys. Chem. Chem. Phys., 1988. **37**(2): p. 785-789.
51. Gill, P.M.W., *A new gradient-corrected exchange functional.* Mol. Phys., 1996. **89** p. 433–445.
52. Becke, A.D., *Density-functional thermochemistry 0.4. A new dynamical correlation functional and implications for exact exchange mixing.* J. Chem. Phys, 1996. **104**: p. 1040–1046.
53. Zhao, Y.T., D. G., *Hybrid meta density functional theory methods for thermochemistry, thermochemical kinetics, and noncovalent interactions: The MPW1B95 and MPWB1K models and comparative assessments for hydrogen bonding and van der Waals interactions.* J. Phys. Chem. A, 2004. **108**: p. 6908–6918.
54. Adamo, C.B., V. , *Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The mPW and mPW1PW models.* J. Chem. Phys., 1998. **108**: p. 664–675.
55. Peverati, R. and D.G. Truhlar, *Exchange-Correlation Functional with Good Accuracy for Both Structural and Energetic Properties while Depending Only on the Density and Its Gradient.* J Chem Theory Comput, 2012. **8**(7): p. 2310-9.
56. Boese, A.D. and N.C. Handy, *A new parametrization of exchange–correlation generalized gradient approximation functionals.* The Journal of Chemical Physics, 2001. **114**(13): p. 5497.
57. Handy, N.C. and A.J. Cohen, *Left-right correlation energy.* Mol. Phys., 2001. **99**(5): p. 403-412.

58. Tao, J., J.P. Perdew, V.N. Staroverov, and G.E. Scuseria, *Climbing the density functional ladder: nonempirical meta-generalized gradient approximation designed for molecules and solids.* Phys Rev Lett, 2003. **91**(14): p. 146401.
59. Perdew, J.P. and A. Zunger, *Self-interaction correction to density-functional approximations for many-electron systems.* Physical Review B, 1981. **23**(10): p. 5048-5079.
60. Goerigk, L. and S. Grimme, *A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions.* Phys Chem Chem Phys, 2011. **13**(14): p. 6670-88.
61. Peverati, R. and D.G. Truhlar, *Improving the Accuracy of Hybrid Meta-GGA Density Functionals by Range Separation.* The Journal of Physical Chemistry Letters, 2011. **2**(21): p. 2810-2817.
62. Becke, A.D., *Density-Functional Thermochemistry 0.3. The Role of Exact Exchange.* J. Chem. Phys. , 1993. **98**: p. 5648–5652.
63. Peverati, R. and D.G. Truhlar, *Screened-exchange density functionals with broad accuracy for chemistry and solid-state physics.* Phys Chem Chem Phys, 2012. **14**(47): p. 16187-91.
64. Van Voorhis, T. and G.E. Scuseria, *A novel form for the exchange-correlation energy functional.* The Journal of Chemical Physics, 1998. **109**(2): p. 400.
65. Becke, A.D., *A new mixing of Hartree–Fock and local density-functional theories.* The Journal of Chemical Physics, 1993. **98**(2): p. 1372.
66. Peverati, R. and D.G. Truhlar, *An improved and broadly accurate local approximation to the exchange-correlation density functional: the MN12-L functional for electronic structure calculations in chemistry and physics.* Phys Chem Chem Phys, 2012. **14**(38): p. 13171-4.
67. Peverati, R. and D.G. Truhlar, *M11-L: A Local Density Functional That Provides Improved Accuracy for Electronic Structure Calculations in Chemistry and Physics.* The Journal of Physical Chemistry Letters, 2012. **3**(1): p. 117-124.
68. Lynch, B.J., P.L. Fast, M. Harris, and D.G. Truhlar, *Adiabatic Connection for Kinetics.* J. Phys. Chem. A, 2000. **104**(21).
69. Kormos, B.L. and C.J. Cramer, *Adiabatic connection method for $X^- + RX$ nucleophilic substitution reactions (X = F, Cl).* J. Phys. Org. Chem., 2002. **15**(10): p. 712-720.
70. Handy, N.C. and A.J. Cohen, *Left-right correlation energy.* Molecular Physics, 2001. **99**(5): p. 403-412.

71. Adamo, C., M. Cossi, and V. Barone, *An accurate density functional method for the study of magnetic properties: the PBE0 model.* Theochem, 1999. **493**: p. 145-157.
72. Hamprecht, F.A., A.J. Cohen, D.J. Tozer, and N.C. Handya, *Development and assessment of new exchange-correlation functionals.* J. Chem. Phys., 1998. **109**(15).
73. Chai, J.D. and M. Head-Gordon, *Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections.* Phys Chem Chem Phys, 2008. **10**(44): p. 6615-20.
74. Wilson, P.J., T.J. Bradley, and D.J. Tozer, *Hybrid exchange-correlation functional determined from thermochemical data and ab initio potentials.* The Journal of Chemical Physics, 2001. **115**(20): p. 9233.
75. Mladek, A., M. Krepl, D. Svozil, P. Cech, M. Otyepka, P. Banas, M. Zgarbova, P. Jurecka, and J. Sponer, *Benchmark quantum-chemical calculations on a complete set of rotameric families of the DNA sugar-phosphate backbone and their comparison with modern density functional theory.* Phys Chem Chem Phys, 2013. **15**(19): p. 7295-310.
76. Jurecka, P., J. Sponer, J. Cerny, and P. Hobza, *Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs.* Phys Chem Chem Phys, 2006. **8**(17): p. 1985-93.
77. Zhao, Y. and D.G. Truhlar, *Density Functionals with Broad Applicability in Chemistry.* Accounts of Chemical Research, 2008. **41**(2).
78. Luo, S., B. Averkiev, K.R. Yang, X. Xu, and D.G. Truhlar, *Density Functional Theory of Open-Shell Systems. The 3d-Series Transition-Metal Atoms and Their Cations.* J Chem Theory Comput, 2014. **10**(1): p. 102-21.
79. Yang, Y., M.N. Weaver, and K.M. Merz, Jr., *Assessment of the "6-31+G** + LANL2DZ" mixed basis set coupled with density functional theory methods and the effective core potential: prediction of heats of formation and ionization potentials for first-row-transition-metal complexes.* J Phys Chem A, 2009. **113**(36): p. 9843-51.
80. Lee, C.-K., C.-W. Pao, and C.-W. Chu, *Multiscale molecular simulations of the nanoscale morphologies of P3HT:PCBM blends for bulk heterojunction organic photovoltaic cells.* Energy & Environmental Science, 2011. **4**(10): p. 4124.

81. Matsumoto, F., T. Iwai, K. Moriwaki, Y. Takao, T. Ito, T. Mizuno, and T. Ohno, *Controlling the Polarity of Fullerene Derivatives to Optimize Nanomorphology in Blend Films.* ACS Appl Mater Interfaces, 2016. **8**(7): p. 4803-10.
82. Goetzbergera, A., C. Heblinga, and H.W. Schockb, *Photovoltaic materials, history, status and outlook.* Materials Science and Engineering R 2002. **40**: p. 1-46.
83. Nelson, J., *Polymer:fullerene bulk heterojunction solar cells.* Materials Today, 2011. **14**(10): p. 462-470.
84. Guo, J., Y. Liang, J. Szarko, B. Lee, H.J. Son, B.S. Rolczynski, L. Yu, and L.X. Chen, *Structure, Dynamics, and Power Conversion Efficiency Correlations in a New Low Bandgap Polymer: PCBM Solar Cell.* J. Phys. Chem. B, 2010. **114**: p. 742-748.
85. Lou, S.J., J.M. Szarko, T. Xu, L. Yu, T.J. Marks, and L.X. Chen, *Effects of additives on the morphology of solution phase aggregates formed by active layer components of high-efficiency organic solar cells.* J Am Chem Soc, 2011. **133**(51): p. 20661-3.
86. Hoppe, H. and N.S. Sariciftci, *Organic solar cells: An overview.* Journal of Materials Research, 2011. **19**(07): p. 1924-1945.
87. Hiramoto, M., H. Fujiwara, and M. Yokoyama, *Three-layered organic solar cell with a photoactive interlayer of codeposited pigments.* Applied Physics Letters, 1991. **58**(10): p. 1062.
88. Yu, G., J. Gao, J.C. Hummelen, F. Wudl, and A.J. Heeger, *Polymer Photovoltaic Cells: Enhanced Efficiencies vis s Network of Internal Donor-Acceptor Heterojunctions.* Science, 1995. **270**.
89. Hou, J. and X. Guo, *Active Layer Materials for Organic Solar Cells.* 2013: p. 17-42.
90. Yonezawa, K., M. Ito, H. Kamioka, T. Yasuda, L. Han, and Y. Moritomo, *Carrier Formation Dynamics of Organic Photovoltaics as Investigated by Time-Resolved Spectroscopy.* Advances in Optical Technologies, 2012. **2012**: p. 1-10.
91. de Gier, H.D., R. Broer, and R.W.A. Havenith, *Non-innocent side-chains with dipole moments in organic solar cells improve charge separation.* Phys. Chem. Chem. Phys, 2014. **16**: p. 12454.
92. Gregg, B.A. and M.C. Hanna, *Comparing organic to inorganic photovoltaic cells: Theory, experiment, and simulation.* Journal of Applied Physics, 2003. **93**(6): p. 3605.

93. Wodo, O., S. Tirthapura, S. Chaudhary, and B. Ganapathysubramanian, *A graph-based formulation for computational characterization of bulk heterojunction morphology.* Organic Electronics, 2012. **13**(6): p. 1105-1113.
94. Razzell-Hollis, J., J. Wade, W.C. Tsoi, Y. Soon, J. Durrant, and J.-S. Kim, *Photochemical stability of high efficiency PTB7:PC70BM solar cell blends.* J. Mater. Chem. A, 2014. **2**(47): p. 20189-20195.
95. Wodo, O., J.D. Roehling, A.J. Moulé, and B. Ganapathysubramanian, *Quantifying organic solar cell morphology: a computational study of three-dimensional maps.* Energy & Environmental Science, 2013. **6**(10): p. 3060.
96. Lindahl, E.R., *Methods in Molecular Biology*, in *Molecular Modelling of Protein*, A. Kukol, Editor. 2008, Humana Press, Totowa, NJ.
97. Marrink, S.J., A.H. de Vries, and A.E. Mark, *Coarse Grained Model for Semiquantitative Lipid Simulations.* J. Phys. Chem. B, 2004. **108**: p. 750-760.
98. Marrink, S.J., E.R. Lindahl, O. Edholm, and A.E. Mark, *Simulation of the Spontaneous Aggregation of Phospholipids into Bilayers.* J. Am. Chem. Soc., 2001. **123**: p. 8638-8639.
99. Muller, M., K. Katsov, and M. Schick, *Coarse-Grained Models and Collective Phenomena in Membranes_ Computer Simulation of Membrane Fusion.* J. Polym. Sci. Part B: Polym. Phys, 2003. **41**.
100. Groot, R.D., T.J. Madden, and D.J. Tildesley, *On the role of hydrodynamic interactions in block copolymer microphase separation.* The Journal of Chemical Physics, 1999. **110**(19): p. 9739.
101. Marrink, S.J., H.J. Risselada, S. Yefimov, D.P. Tieleman, and A.H. de Vries, *The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations.* J. Phys. Chem. B, 2007. **111**: p. 7812-7824.
102. Marrink, S.J. and D.P. Tieleman, *Perspective on the Martini model.* Chem Soc Rev, 2013. **42**(16): p. 6801-22.
103. de Jong, D.H., G. Singh, W.F. Bennett, C. Arnarez, T.A. Wassenaar, L.V. Schafer, X. Periole, D.P. Tieleman, and S.J. Marrink, *Improved Parameters for the Martini Coarse-Grained Protein Force Field.* J Chem Theory Comput, 2013. **9**(1): p. 687-97.
104. Monticelli, L., S.K. Kandasamy, X. Periole, R.G. Larson, D.P. Tieleman, and S.J. Marrink, *The MARTINI Coarse-Grained Force Field: Extension to Proteins.* J. Chem. Theory and Comput., 2008. **4**(819-834).
105. Monticelli, L., *On Atomistic and Coarse-Grained Models for C60 Fullerene.* J Chem Theory Comput, 2012. **8**(4): p. 1370-8.

106. Monticelli, L., *http://perso.ibcp.fr/luca.monticelli/MARTINI/index.html.*
107. Alessandri, R., J.J. Uusitalo, A.H. de Vries, R.W.A. Havenith, and S.J. Marrink, *soon to be published.* 2016.
108. Malde, A.K., L. Zuo, M. Breeze, M. Stroet, D. Poger, P.C. Nair, C. Oostenbrink, and A.E. Mark, *An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0.* J Chem Theory Comput, 2011. **7**(12): p. 4026-37.
109. Oostenbrink, C., A. Villa, A.E. Mark, and W.F. van Gunsteren, *A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6.* J Comput Chem, 2004. **25**(13): p. 1656-76.
110. de Jong, D.H., S. Baoukina, H.I. Ingólfsson, and S.J. Marrink, *Martini straight: Boosting performance using a shorter cutoff and GPUs.* Computer Physics Communications, 2016. **199**: p. 1-7.
111. Abraham, M.J., T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, and E. Lindahl, *GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers.* SoftwareX, 2015. **1-2**: p. 19-25.
112. Bussi, G., D. Donadio, and M. Parrinello, *Canonical sampling through velocity rescaling.* J Chem Phys, 2007. **126**(1): p. 014101.
113. Parrinello, M. and A. Rahman, *Polymorphic transitions in single crystals: A new molecular dynamics method.* Journal of Applied Physics, 1981. **52**(12): p. 7182.

Supporting information

I. Benchmarking of Density Functionals for the Electron Affinity of the Fe(III)/Fe(II) Redox Pair

1. Optimisation energy of Fe (II) and Fe (III) complexes

		Optimisation Energy (Hartree)		
	Model	Fe (II) high spin	Fe (II) low spin	Δ energy (kcal/mol)
1	$Fe(H_2O)$	-1338.076338	-1337.910738	-103.92
2	$Fe(H_2O)_2$	-1414.466654	-1414.319378	-92.42
3	$Fe(H_2O)_4$	-1567.184584	-1567.056098	-80.63
4	$Fe(H_2O)_6$	-1719.8341	-1719.745572	-55.55
5	$Fe(OCH_3)$	-1377.127358	-1376.978548	-93.38
6	Fe(OOCH)	-1451.021573	-1450.927924	-58.77
7	$Fe(SCH_3)$	-1699.724901	-1699.565244	-100.19
8	$Fe(C_6H_5O)$	-1568.352644	-1568.233159	-74.98
9	$Fe(H_2O)_3(NH_2CH_3)$	-1586.531863	-1586.410118	-76.40
10	$Fe(H_2O)_3(COOH)$	-1680.057524		
11	$Fe(H_2O)_3(OCH_3)$	-1606.121691	-1606.004651	-73.44
12	$Fe(H_2O)_3(SCH_3)$	-1928.718495	-1928.60493	-71.26
13	$Fe(H_2O)_5(NH_2CH_3)$	-1739.170501	-1739.086065	-52.98
14	$Fe(H_2O)_5(COOH)$	-1832.681779	-1832.602509	-49.74
15	$Fe(H_2O)_5(OCH_3)$	-1758.735127	-1758.653081	-51.48
16	$Fe(H_2O)_5(SCH_3)$	-2081.330621	-2081.250132	-50.51
		MSE		-72.38

		Optimisation Energy (Hartree)		
	Model	Fe (III) high spin	Fe (III) low spin	Δ energy (kcal/mol)
1	$Fe(H_2O)$	-1337.127819	-1336.946646	-113.69
2	$Fe(H_2O)_2$	-1413.618236	-1413.473901	-90.57
3	$Fe(H_2O)_4$	-1566.485897	-1566.349327	-85.70
4	$Fe(H_2O)_6$	-1719.22353	-1719.106471	-73.46
5	$Fe(OCH_3)$	-1376.579978	-	-
6	Fe(OOCH)	-1450.48914	-	-
7	$Fe(SCH_3)$	-1699.222572	-1699.152289	-44.10
8	$Fe(C_6H_5O)$	-1567.907522	-1567.812328	-59.73
9	$Fe(H_2O)_3(NH_2CH_3)$	-1585.861735	-1585.751177	-69.38
10	$Fe(H_2O)_3(COOH)$	-1679.578715	-	-
11	$Fe(H_2O)_3(OCH_3)$	-1605.691299	-1605.598838	-58.02
12	$Fe(H_2O)_3(SCH_3)$	-1928.311944	-1928.239246	-45.62
13	$Fe(H_2O)_5(NH_2CH_3)$	-1738.57739	-	-

14	$Fe(H_2O)_5(COOH)$	-1832.251353	-1832.159822	-57.44
15	$Fe(H_2O)_5(OCH_3)$	-1758.346618	-1758.240454	-66.62
16	$Fe(H_2O)_5(SCH_3)$	-2080.956763	-2080.88633	-44.20
		MSE		-67.38

2. Single point energy at MP2/aug-cc-pVXZ level (X = 2, 3, 4) and CCSD(T)/aug-cc-pVDZ

		Fe (II) (Hartree)			
	Model	SP with MP2/aug-cc-pVDZ	SP with MP2/aug-cc-pVTZ	SP with MP2/aug-cc-pVQZ	CCSD(T)/ aug-cc-pVDZ
1	$Fe(H_2O)$	-1338.152552	-1338.25058	-1338.286357	-1338.1723
2	$Fe(H_2O)_2$	-1414.529124	-1414.695187	-1414.754278	-1414.5631
3	$Fe(H_2O)_4$	-1567.216503	-1567.518087	-1567.623441	-1567.2782
4	$Fe(H_2O)_6$	-1719.836083	-1720.270264	-1720.420675	-1719.9245
5	$Fe(OCH_3)$	-1377.201995	-1377.339737	-1377.387079	-1377.2478
6	Fe(OOCH)	-1451.083743	-1451.274089	-1451.340938	-1451.1247
7	$Fe(SCH_3)$	-1699.815766	-1699.939235	-1699.981882	-1699.8756
8	$Fe(C_6H_5O)$	-1568.397566	-1568.692287	-1568.791095	-1568.4934
9	$Fe(H_2O)_3(NH_2CH_3)$	-1586.559918	-1586.885617	-1586.996309	-1586.6492
10	$Fe(H_2O)_3(COOH)$	-1680.078643	-1680.470973	-1680.606821	-1680.1652
11	$Fe(H_2O)_3(OCH_3)$	-1606.15196	-1606.491213	-1606.60756	-1606.2350
12	$Fe(H_2O)_3(SCH_3)$	-1928.767371	-1929.093364	-1929.205387	-1928.8622
13	$Fe(H_2O)_5(NH_2CH_3)$	-1739.168776	-1739.627106	-1739.782734	-1739.2847
14	$Fe(H_2O)_5(COOH)$	-1832.673315	-1833.198132	-1833.378929	-1832.7765
15	$Fe(H_2O)_5(OCH_3)$	-1758.736541	-1759.208556	-1759.3698	-1758.8455
16	$Fe(H_2O)_5(SCH_3)$	-2081.346981	-2081.805566	-2081.962415	-2081.4662

		Fe (III) (Hartree)			
	Model	SP with MP2/aug-cc-pVDZ	SP with MP2/aug-cc-pVTZ	SP with MP2/aug-cc-pVQZ	CCSD(T)/ aug-cc-pVDZ
1	$Fe(H_2O)$	-1337.201722	-1337.295811	-1337.327394	-1337.235939
2	$Fe(H_2O)_2$	-1413.680864	-1413.843157	-1413.898248	-1413.724433
3	$Fe(H_2O)_4$	-1566.520824	-1566.818645	-1566.920278	-1566.589275
4	$Fe(H_2O)_6$	-1719.228374	-1719.659049	-1719.805715	-1719.322043
5	$Fe(OCH_3)$	-1376.651543	-1376.780085	-1376.823898	-1376.700891
6	Fe(OOCH)	-1450.546942	-1450.730839	-1450.794895	-1450.593894
7	$Fe(SCH_3)$	-1699.310794	-1699.429252	-1699.469416	-1699.366529
8	$Fe(C_6H_5O)$	-1567.931094	-1568.221276	-1568.318522	-1568.034824
9	$Fe(H_2O)_3(NH_2CH_3)$	-1585.89264	-1586.213676	-1586.32063	-1585.990101
10	$Fe(H_2O)_3(COOH)$	-1679.599311	-1679.987367	-1680.119879	-1679.69183
11	$Fe(H_2O)_3(OCH_3)$	-1605.718763	-1606.053115	-1606.166239	-1605.815978
12	$Fe(H_2O)_3(SCH_3)$	-1928.354961	-1928.67704	-1928.786698	-1928.450809
13	$Fe(H_2O)_5(NH_2CH_3)$	-1738.578472	-1739.03251	-1739.184349	-1738.700116
14	$Fe(H_2O)_5(COOH)$	-1832.244511	-1832.765674	-1832.943162	-1832.357513
15	$Fe(H_2O)_5(OCH_3)$	-1758.344388	-1758.812224	-1758.97034	-1758.467981
16	$Fe(H_2O)_5(SCH_3)$	-2080.969968	-2081.424757	-2081.579203	-2081.09239

3. Benchmarking of DFT

Group A (Fe complexes coordinated with water molecules as ligands)

FUNCTIONAL	$[Fe(H_2O)]^{2+}$	$[Fe(H_2O)]^{3+}$	ΔEel (kcal/mol)	ΔDeltaE (kcal/mol)	$[Fe(H_2O)_2]^{2+}$	$[Fe(H_2O)_2]^{3+}$	ΔEel (kcal/mol)	ΔDeltaE (kcal/mol)
CCSD(T)/CBS	-1338.326846	-1337.3814	-593.2473174	0.00	-1414.827476	-1413.977277	-533.5079098	0.00
MPWB1K	-1339.469989	-1338.523	-594.2224183	0.98	-1416.008036	-1415.157799	-533.5322061	0.02
BB1K	-1339.464139	-1338.5186	-593.3634455	0.12	-1416.002875	-1415.15435	-532.4574327	1.05
MPW1N	-1339.379224	-1338.431	-595.0362104	1.79	-1415.927812	-1415.075557	-534.7982255	1.29
MPW1K	-1339.375593	-1338.4266	-595.522066	2.27	-1415.923014	-1415.06957	-535.5440395	2.04
M062X	-1339.27606	-1338.3226	-598.3307362	5.08	-1415.823605	-1414.969371	-536.0400043	2.53
BMK	-1338.759689	-1337.808	-597.2180173	3.97	-1415.29966	-1414.449045	-533.7694863	0.26
MN12SX	-1339.163484	-1338.2159	-594.6217842	1.37	-1415.676814	-1414.830894	-530.8226134	2.69
MPW1B95	-1339.492881	-1338.5504	-591.4056281	1.84	-1416.038244	-1415.19661	-528.1333143	5.37
B3LYP	-1339.359924	-1338.4147	-593.143792	0.10	-1415.941907	-1415.102339	-526.836779	6.67
B3PW91	-1339.285747	-1338.3431	-591.5148901	1.73	-1415.836037	-1414.997802	-526.0003339	7.51
M11L	-1339.41487	-1338.4672	-594.6767164	1.43	-1415.962565	-1415.12281	-526.9543743	6.55
M052X	-1339.265552	-1338.3076	-601.1086333	7.86	-1415.83183	-1414.974218	-538.1596376	4.65
MPW2PLYP	-1338.914325	-1337.9737	-590.2387175	3.01	-1415.359097	-1414.513915	-530.3598187	3.15
B1B95	-1339.486944	-1338.5464	-590.2295119	3.02	-1416.033562	-1415.195012	-526.1985265	7.31
B1LYP	-1339.319896	-1338.3797	-590.0075053	3.24	-1415.869205	-1415.031704	-525.540294	7.97
OVWN5	-1340.921758	-1339.9738	-594.8484783	1.60	-1417.794661	-1416.952272	-528.6072158	4.90
OPL	-1340.897866	-1339.9485	-595.723051	2.48	-1417.76652	-1416.923143	-529.2272016	4.28
MN12L	-1338.978701	-1338.0383	-590.0916544	3.16	-1415.47207	-1414.63619	-524.5228374	8.99
PBE1PBE	-1339.029389	-1338.0892	-589.9856554	3.26	-1415.525368	-1414.687153	-525.9882104	7.52
B2PLYP	-1338.880866	-1337.9422	-589.019435	4.23	-1415.319289	-1414.475888	-529.2419543	4.27
B2GPPLYP	-1338.73313	-1337.7944	-589.0591563	4.19	-1415.133967	-1414.292459	-528.0540347	5.45

FUNCTIONAL	$[Fe(H_2O)]^{2+}$	$[Fe(H_2O)]^{3+}$	ΔEel (kcal/mol)	ΔDeltaE (kcal/mol)	$[Fe(H_2O)_2]^{2+}$	$[Fe(H_2O)_2]^{3+}$	ΔEel (kcal/mol)	ΔDeltaE (kcal/mol)
wB97X-D	-1339.333967	-1338.3854	-595.207282	1.96	-1415.857831	-1415.039843	-513.2949525	20.21
M11	-1339.217951	-1338.2583	-602.1679699	8.92	-1415.769042	-1414.907722	-540.4863051	6.98
DSD_BLYP	-1338.610019	-1337.6732	-587.8445299	5.40	-1414.98036	-1414.14232	-525.877982	7.63
TPSS	-1339.345962	-1338.4121	-586.0266283	7.22	-1415.920261	-1415.09267	-519.3214731	14.19
TPSSTPSS	-1339.345962	-1338.4121	-586.0266283	7.22	-1415.920261	-1415.09267	-519.3214668	14.19
BHandH	-1337.082065	-1336.1434	-589.0223529	4.22	-1413.197065	-1412.357465	-526.8574555	6.65
HCTH407	-1339.854603	-1338.9123	-591.3317639	1.92	-1416.42636	-1415.594558	-521.9638972	11.54
SVWN	-1336.999248	-1336.0401	-601.873988	8.63	-1413.237572	-1412.398437	-526.5648352	6.94
B3P86	-1340.048262	-1339.0832	-605.5602304	12.31	-1416.807819	-1415.947846	-539.6410058	6.13
N12	-1339.720452	-1338.7763	-592.4466289	0.80	-1416.257311	-1415.433067	-517.2208157	16.29
BP86	-1339.501595	-1338.5624	-589.3778183	3.87	-1416.086565	-1415.264217	-516.0314653	17.48
M06	-1339.226625	-1338.2953	-584.4214588	8.83	-1415.766938	-1414.939862	-518.9981174	14.51
B97-1	-1339.19841	-1338.2764	-578.5859082	14.66	-1415.7482	-1414.934095	-510.8587343	22.65
B97-2	-1339.536175	-1338.6162	-577.3082547	15.94	-1416.080304	-1415.268063	-509.6889372	23.82
M05	-1339.307642	-1338.3798	-582.2432412	11.00	-1415.844917	-1415.0219	-516.4510371	17.06
B97D3	-1339.767275	-1338.8311	-587.4425473	5.80	-1416.311354	-1415.484994	-518.5486009	14.96
BPW91	-1339.444635	-1338.5097	-586.6580222	6.59	-1416.017619	-1415.198905	-513.7507692	19.76
MPWB95	-1339.56418	-1338.626	-588.7394527	4.51	-1416.131973	-1415.310379	-515.5578085	17.95
BPBE	-1339.322203	-1338.3886	-585.8225685	7.42	-1415.870881	-1415.053349	-513.0092285	20.50
OLYP	-1339.608236	-1338.6776	-583.9821581	9.27	-1416.160633	-1415.341851	-513.7937096	19.71
G96LYP	-1339.474623	-1338.542	-585.2535553	7.99	-1416.038044	-1415.222885	-511.5204494	21.99
VSXC	-1339.847241	-1338.9258	-578.2211619	15.03	-1416.44591	-1415.632643	-510.3326742	23.18
OTPSS-D	-1339.561066	-1338.6363	-580.2878334	12.96	-1416.095293	-1415.284054	-509.0600219	24.45
M06L	-1339.269927	-1338.3516	-576.2867255	16.96	-1415.832292	-1415.024142	-507.1223978	26.39
SPW91	-1335.059217	-1334.1444	-574.0305461	19.22	-1410.795141	-1409.999581	-499.2213805	34.29

FUNCTIONAL	$[Fe(H_2O)_4]^{2+}$	$[Fe(H_2O)_4]^{3+}$	ΔEel (kcal/mol)	ΔDeltaE (kcal/mol)	$[Fe(H_2O)_6]^{2+}$	$[Fe(H_2O)_6]^{3+}$	ΔEel (kcal/mol)	ΔDeltaE (kcal/mol)	MUE (kcal/mol)	MaxE (kcal/mol)
CCSD(T)/CBS	-1567.754505	-1567.05455	-439.228576	0.00	-1720.608047	-1719.994729	-384.862995	0.00	0.00	0.00
MPWB1K	-1569.013235	-1568.314961	-438.1732933	1.06	-1721.95166	-1721.337742	-385.2393884	0.38	0.61	1.06
BB1K	-1569.009062	-1568.311665	-437.6230803	1.61	-1721.947739	-1721.334099	-385.0648717	0.20	0.74	1.61
MPW1N	-1568.953206	-1568.252724	-439.5595122	0.33	-1721.909086	-1721.292693	-386.79287	1.93	1.34	1.93
MPW1K	-1568.946683	-1568.24571	-439.8674939	0.64	-1721.901094	-1721.284661	-386.8177508	1.95	1.73	2.27
M062X	-1568.853674	-1568.154438	-438.7771081	0.45	-1721.811456	-1721.199145	-384.2314511	0.63	2.17	5.08
BMK	-1568.304708	-1567.608645	-436.7862837	2.44	-1721.241843	-1720.625245	-386.9210702	2.06	2.18	3.97
MN12SX	-1568.640997	-1567.939603	-440.1312801	0.90	-1721.542594	-1720.919914	-390.7376275	5.87	2.71	5.87
MPW1B95	-1569.054383	-1568.36065	-435.3240107	3.90	-1722.002124	-1721.389865	-384.1983814	0.66	2.95	5.37
B3LYP	-1569.028062	-1568.334411	-435.272831	3.96	-1722.040602	-1721.424313	-386.7272764	1.86	3.15	6.67
B3PW91	-1568.858486	-1568.166007	-434.5373207	4.69	-1721.807678	-1721.194598	-384.713435	0.15	3.52	7.51
M11L	-1568.991178	-1568.291169	-439.2620915	0.03	-1721.959936	-1721.336871	-390.9795325	6.12	3.53	6.55
M052X	-1568.901356	-1568.199522	-440.4070706	1.18	-1721.899742	-1721.283901	-386.4458572	1.58	3.82	7.86
MPW2PLYP	-1568.181639	-1567.491303	-433.1928939	6.04	-1720.936551	-1720.329208	-381.1136822	3.75	3.99	6.04
B1B95	-1569.05146	-1568.359728	-434.0684142	5.16	-1721.999935	-1721.388815	-383.4839052	1.38	4.22	7.31
B1LYP	-1568.891935	-1568.201135	-433.4833179	5.75	-1721.842155	-1721.229256	-384.5999938	0.26	4.30	7.97
OVWN5	-1571.440675	-1570.745455	-436.2570923	2.97	-1725.0188	-1724.392941	-392.732468	7.87	4.34	7.87
OPL	-1571.403811	-1570.707765	-436.7750262	2.45	-1724.973065	-1724.346326	-393.2847517	8.42	4.41	8.42
MN12L	-1568.393852	-1567.696414	-437.6495612	1.58	-1721.256108	-1720.635342	-389.5364424	4.67	4.60	8.99
PBE1PBE	-1568.441098	-1567.750206	-433.5416073	5.69	-1721.285702	-1720.675495	-382.9110266	1.95	4.61	7.52
B2PLYP	-1568.128802	-1567.439545	-432.5153781	6.71	-1720.870136	-1720.263246	-380.8292007	4.03	4.81	6.71
B2GPPLYP	-1567.871552	-1567.186038	-430.1664653	9.06	-1720.542412	-1719.940878	-377.468297	7.39	6.52	9.06
wB97X-D	-1568.92278	-1568.22635	-437.0160087	2.21	-1721.893846	-1721.276987	-387.084869	2.22	6.65	20.21

FUNCTIONAL	$[Fe(H_2O)_4]^{2+}$	$[Fe(H_2O)_4]^{3+}$	ΔEel (kcal/mol)	ΔDeltaE (kcal/mol)	$[Fe(H_2O)_6]^{2+}$	$[Fe(H_2O)_6]^{3+}$	ΔEel (kcal/mol)	ΔDeltaE (kcal/mol)	MUE (kcal/mol)	MaxE (kcal/mol)
M11	-1568.804862	-1568.094049	-446.0416424	6.81	-1721.782084	-1721.154239	-393.9791413	9.12	7.96	9.12
DSD_BLYP	-1567.658358	-1566.976504	-427.8695417	11.36	-1720.270292	-1719.672752	-374.962168	9.90	8.57	11.36
TPSS	-1568.98951	-1568.305061	-429.4983182	9.73	-1721.984385	-1721.380465	-378.9656855	5.90	9.26	14.19
TPSSTPSS	-1568.98951	-1568.305061	-429.4983182	9.73	-1721.984385	-1721.380465	-378.9656855	5.90	9.26	14.19
BHandH	-1565.351308	-1564.671396	-426.6508867	12.58	-1717.435256	-1716.844968	-370.4114746	14.45	9.48	14.45
HCTH407	-1569.478068	-1568.799982	-425.5053805	13.72	-1722.459827	-1721.829348	-395.6316878	10.77	9.49	13.72
SVWN	-1565.613084	-1564.932435	-427.1135809	12.11	-1717.912169	-1717.31566	-374.3151491	10.55	9.56	12.11
B3P86	-1570.248315	-1569.535037	-447.5887674	8.36	-1723.614802	-1722.981619	-397.3288942	12.47	9.82	12.47
N12	-1569.247526	-1568.570215	-425.018615	14.21	-1722.164569	-1721.565217	-376.0992467	8.76	10.02	16.29
BP86	-1569.170434	-1568.493571	-424.7383943	14.49	-1722.176957	-1721.576782	-376.6151224	8.25	11.02	17.48
M06	-1568.784517	-1568.105501	-426.0891339	13.14	-1721.737127	-1721.136035	-377.1913331	7.67	11.04	14.51
B97-1	-1568.804717	-1568.116404	-431.9227329	7.31	-1721.767101	-1721.156333	-383.2626265	1.60	11.55	22.65
B97-2	-1569.129756	-1568.441834	-431.6778159	7.55	-1722.082644	-1721.472523	-382.8568537	2.01	12.33	23.82
M05	-1568.856127	-1568.179628	-424.5099055	14.72	-1721.794727	-1721.192621	-377.8273956	7.04	12.45	17.06
B97D3	-1569.313319	-1568.64268	-420.8319714	18.40	-1722.24793	-1721.653187	-373.2070675	11.66	12.70	18.40
BPW91	-1569.078236	-1568.403873	-423.1691309	16.06	-1722.062516	-1721.4645	-375.2609818	9.60	13.00	19.76
MPWB95	-1569.172349	-1568.502109	-420.5823292	18.65	-1722.144935	-1721.55116	-372.599381	12.26	13.34	18.65
BPBE	-1568.882794	-1568.209576	-422.450852	16.78	-1721.818356	-1721.22167	-374.4260363	10.44	13.78	20.50
OLYP	-1569.173306	-1568.503899	-420.0591557	19.17	-1722.117033	-1721.520206	-374.5144963	10.35	14.62	19.71
G96LYP	-1569.078894	-1568.408565	-420.6380333	18.59	-1722.040797	-1721.445278	-373.6937326	11.17	14.94	21.99
VSXC	-1569.5753	-1568.903489	-421.5679397	17.66	-1722.671367	-1722.051376	-389.0502417	4.19	15.01	23.18
OTPSS-D	-1569.08748	-1568.416551	-421.0143068	18.21	-1722.003408	-1721.406447	-374.5988023	10.26	16.47	24.45
M06L	-1568.893021	-1568.229124	-416.6016091	22.63	-1721.886096	-1721.299601	-368.0317574	16.83	20.70	26.39
SPW91	-1562.167836	-1561.528767	-401.0221585	38.21	-1713.462171	-1712.912227	-345.0950893	39.77	32.87	39.77

Group B (Fe complexes with a single ligand that represents amino acid sidechain)

FUNCTIONAL	$[Fe(OCH_3)]^+$	$[Fe(OCH_3)]^+$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	$[Fe(OOCH)]^+$	$Fe(OOCH)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)
CCSD(T)/CBS	-1377.461344	-1376.900312	-352.052534	0.00	-1451.425807	-1450.883631	-340.2207951	0.00
PBE1PBE	-1378.196736	-1377.636233	-351.7212957	0.33	-1452.165086	-1451.622598	-340.4161581	-0.20
MPW1B95	-1378.683553	-1378.122695	-351.9435596	0.11	-1452.711188	-1452.169609	-339.8457142	0.38
B1B95	-1378.679829	-1378.120699	-350.8595305	1.19	-1452.70891	-1452.166384	-340.4398967	-0.22
B1LYP	-1378.516494	-1377.956677	-351.290617	0.76	-1452.543859	-1452.003128	-339.3134041	0.91
B3PW91	-1378.49387	-1377.932685	-352.1492447	-0.10	-1452.500771	-1451.951321	-344.7846854	-4.56
BMK	-1377.93917	-1377.366921	-359.091898	-7.04	-1451.95733	-1451.418237	-338.2864783	1.93
wB97X-D	-1378.53089	-1377.964991	-355.1069736	-3.05	-1452.544386	-1452.001315	-340.7823852	-0.56
MPWB1K	-1378.648083	-1378.088682	-351.029184	1.02	-1452.666116	-1452.135032	-333.260572	6.96
B3LYP	-1378.589914	-1378.025739	-354.0248395	-1.97	-1452.62803	-1452.075633	-346.6348786	-6.41
BB1K	-1378.643235	-1378.076683	-355.5166181	-3.46	-1452.662006	-1452.130694	-333.4038889	6.82
B97-1	-1378.424187	-1377.865632	-350.4987752	1.55	-1452.423355	-1451.880598	-340.5851652	-0.36
MN12L	-1378.154915	-1377.594961	-351.376109	0.68	-1452.155356	-1451.611241	-341.4369279	-1.22
M06	-1378.414635	-1377.856512	-350.2272581	1.83	-1452.429173	-1451.878834	-345.3427045	-5.12
MPW1N	-1378.571198	-1378.000607	-358.0513302	-6.00	-1452.578548	-1452.04834	-332.7105159	7.51
M062X	-1378.459426	-1377.879573	-363.8633115	-11.81	-1452.484322	-1451.928024	-349.0819652	-8.86
TPSS	-1378.582741	-1378.028177	-347.9945789	4.06	-1452.619546	-1452.07026	-344.6826209	-4.46
TPSSTPSS	-1378.582741	-1378.028177	-347.9945789	4.06	-1452.619546	-1452.07026	-344.6826209	-4.46
M11	-1378.393857	-1377.813299	-364.3058501	-12.25	-1452.41964	-1451.880486	-338.324223	1.90
HCTH407	-1379.087898	-1378.533198	-348.0800771	3.97	-1453.075989	-1452.531087	-341.9315435	-1.71
M05	-1378.493293	-1377.941368	-346.3377967	5.71	-1452.51459	-1451.969385	-342.1212271	-1.90
M11L	-1378.632999	-1378.066607	-355.4159216	-3.36	-1452.649754	-1452.095541	-347.7741976	-7.55
MN12SX	-1378.342005	-1377.773572	-356.6973464	-4.64	-1452.343055	-1451.784851	-350.2784127	-10.06

FUNCTIONAL	$[Fe(OCH_3)]^+$	$[Fe(OCH_3)]^+$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	$[Fe(OOCH)]^+$	$Fe(OOCH)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)
MPW1K	-1378.565673	-1377.993757	-358.8824039	-6.83	-1452.571459	-1452.042971	-331.6310174	8.59
M052X	-1378.463225	-1377.878819	-366.7201055	-14.67	-1452.500253	-1451.9441	-348.9915787	-8.77
B97-2	-1378.733352	-1378.201162	-333.9541458	18.10	-1452.746014	-1452.205618	-339.1036088	1.12
M06L	-1378.491018	-1377.949338	-339.90999	12.14	-1452.527259	-1451.979065	-343.99728	-3.78
VSXC	-1379.088474	-1378.550074	-337.8511247	14.20	-1453.148926	-1452.59843	-345.4411796	-5.22
OTPSS-D	-1378.77214	-1378.230923	-339.6188695	12.43	-1452.780767	-1452.224439	-349.1014996	-8.88
BPBE	-1378.542155	-1377.990987	-345.8634057	6.19	-1452.562837	-1451.998082	-354.3891465	-14.17
N12	-1378.965938	-1378.411575	-347.8680228	4.18	-1452.994206	-1452.431663	-353.0008756	-12.78
BPW91	-1378.681207	-1378.128872	-346.5952576	5.46	-1452.72197	-1452.156308	-354.9587558	-14.74
B97D3	-1378.974744	-1378.425209	-344.8380802	7.21	-1452.973173	-1452.412924	-351.5614627	-11.34
BHandH	-1375.970814	-1375.403265	-356.1421071	-4.09	-1449.620882	-1449.095633	-329.5984887	10.62
G96LYP	-1378.695703	-1378.14608	-344.8936712	7.16	-1452.741479	-1452.178944	-352.9962634	-12.78
OLYP	-1378.821155	-1378.27851	-340.5152418	11.54	-1452.855913	-1452.299381	-349.2289468	-9.01
OVWN5	-1380.371653	-1379.810545	-352.1002801	-0.05	-1454.621714	-1454.049945	-358.7902792	-18.57
BP86	-1378.748581	-1378.190836	-349.9903795	2.06	-1452.794247	-1452.222637	-358.691026	-18.47
OPL	-1380.346548	-1379.784193	-352.883569	-0.83	-1454.592152	-1454.018857	-359.7480785	-19.53
SPW91	-1373.73204	-1373.190304	-339.9444465	12.11	-1447.055097	-1446.486546	-356.7706647	-16.55
B2PLYP	-1377.979916	-1377.415275	-354.3179995	-2.27	-1451.920436	-1451.413926	-317.8400193	22.38
MPW2PLYP	-1378.017918	-1377.450274	-356.20189	-4.15	-1451.962513	-1451.45546	-318.1803742	22.04
B3P86	-1379.420091	-1378.836206	-366.3936624	-14.34	-1453.557264	-1452.984548	-359.3847692	-19.16
MPWB95	-1378.788932	-1378.236891	-346.4111651	5.64	-1452.844843	-1452.275722	-357.1288471	-16.91
B2GPPLYP	-1377.8007	-1377.231338	-357.2801648	-5.23	-1451.711049	-1451.2164	-310.3965067	29.82
DSD_BLYP	-1377.653746	-1377.083969	-357.5405248	-5.49	-1451.541151	-1451.053082	-306.2680143	33.95
SVWN	-1376.043042	-1375.46371	-363.5360653	-11.48	-1449.714035	-1449.111709	-377.9655921	-37.74

FUNCTIONAL	$[Fe(SCH_3)]^+$	$[Fe(SCH_3)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	$[Fe(C_6H_5O)]^+$	$[Fe(C_6H_5O)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	MUE (kcal/mol)	MaxE (kcal/mol)
CCSD(T)/CBS	-1700.070061	-1699.551543	-325.3746436	0.00	-1568.952155	-1568.486151	-292.4218874	0.00	0.00	0.00
PBE1PBE	-1701.107793	-1700.586279	-327.2553678	-1.88	-1569.78288	-1569.319036	-291.0666131	1.36	0.94	1.88
MPW1B95	-1701.733178	-1701.209444	-328.6483011	-3.27	-1570.412074	-1569.948984	-290.5939289	1.83	1.40	3.27
B1B95	-1701.722837	-1701.198687	-328.9095333	-3.53	-1570.410303	-1569.947971	-290.1172915	2.30	1.81	3.53
B1LYP	-1701.504127	-1700.986822	-324.6139415	0.76	-1570.226629	-1569.772923	-284.7045819	7.72	2.54	7.72
B3PW91	-1701.460988	-1700.934324	-330.4868917	-5.11	-1570.226804	-1569.76194	-291.7061458	0.72	2.62	5.11
BMK	-1700.833993	-1700.315521	-325.3466593	0.03	-1569.620567	-1569.159479	-289.3369645	3.08	3.02	7.04
wB97X-D	-1701.521156	-1700.994127	-330.7151546	-5.34	-1570.268106	-1569.807705	-288.9062607	3.52	3.12	5.34
MPWB1K	-1701.702853	-1701.186713	-323.8831187	1.49	-1570.367868	-1569.909908	-287.3745474	5.05	3.63	6.96
B3LYP	-1701.574675	-1701.047412	-330.8623683	-5.49	-1570.398559	-1569.935838	-290.3616061	2.06	3.98	6.41
BB1K	-1701.692694	-1701.176571	-323.8720745	1.50	-1570.364103	-1569.906896	-286.901788	5.52	4.33	6.82
B97-1	-1701.37627	-1700.851372	-329.3781073	-4.00	-1570.167478	-1569.719623	-281.0331927	11.39	4.33	11.39
MN12L	-1701.158495	-1700.625639	-334.3721864	-9.00	-1569.807731	-1569.357253	-282.6795771	9.74	5.16	9.74
M06	-1701.400734	-1700.873695	-330.721756	-5.35	-1570.077699	-1569.626056	-283.4101551	9.01	5.33	9.01
MPW1N	-1701.58743	-1701.073278	-322.6353786	2.74	-1570.326191	-1569.869837	-286.3667733	6.06	5.58	7.51
M062X	-1701.438698	-1700.923045	-323.5773521	1.80	-1570.202291	-1569.736441	-292.324864	0.10	5.64	11.81
TPSS	-1701.568424	-1701.043428	-329.4401366	-4.07	-1570.407654	-1569.958559	-281.8109846	10.61	5.80	10.61
TPSSTPSS	-1701.568424	-1701.043428	-329.4401366	-4.07	-1570.407654	-1569.958559	-281.810972	10.61	5.80	10.61
M11	-1701.368837	-1700.855941	-321.8475152	3.53	-1570.082591	-1569.607156	-298.3403781	-5.92	5.90	12.25
HCTH407	-1702.137959	-1701.595957	-340.1114582	-14.74	-1570.860449	-1570.400276	-288.7631069	3.66	6.02	14.74
M05	-1701.505723	-1700.982486	-328.3359958	-2.96	-1570.144749	-1569.701894	-277.89604	14.53	6.28	14.53
M11L	-1701.619185	-1701.087787	-333.4568695	-8.08	-1570.40748	-1569.952383	-285.5777866	6.84	6.46	8.08
MN12SX	-1701.322648	-1700.803067	-326.0422851	-0.67	-1570.004045	-1569.555587	-281.4122524	11.01	6.59	11.01

FUNCTIONAL	$[Fe(SCH_3)]^+$	$[Fe(SCH_3)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	$[Fe(C_6H_5O)]^+$	$[Fe(C_6H_5O)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	MUE (kcal/mol)	MaxE (kcal/mol)
MPW1K	-1701.582768	-1701.071452	-320.8556359	4.52	-1570.319344	-1569.863911	-285.7889938	6.63	6.64	8.59
M052X	-1701.431205	-1700.91727	-322.4988702	2.88	-1570.266515	-1569.802649	-291.0800104	1.34	6.91	14.67
B97-2	-1701.785233	-1701.256146	-332.007109	-6.63	-1570.499182	-1570.037128	-289.9432956	2.48	7.08	18.10
M06L	-1701.469372	-1700.94277	-330.4478669	-5.07	-1570.279997	-1569.828334	-283.4225924	9.00	7.50	12.14
VSXC	-1702.194632	-1701.668047	-330.4366407	-5.06	-1570.997285	-1570.542445	-285.4161025	7.01	7.87	14.20
OTPSS-D	-1701.76833	-1701.239354	-331.9375307	-6.56	-1570.503652	-1570.047875	-286.0048634	6.42	8.57	12.43
BPBE	-1701.503525	-1700.970929	-334.2092598	-8.83	-1570.2567	-1569.799293	-287.027246	5.39	8.65	14.17
N12	-1701.971912	-1701.433611	-337.7888444	-12.41	-1570.819455	-1570.36414	-285.7144959	6.71	9.02	12.78
BPW91	-1701.678529	-1701.141994	-336.6803675	-11.31	-1570.468974	-1570.011435	-287.110529	5.31	9.20	14.74
B97D3	-1702.007554	-1701.473346	-335.2201338	-9.85	-1570.654889	-1570.202992	-283.5696932	8.85	9.31	11.34
BHandH	-1698.391326	-1697.88361	-318.5964193	6.78	-1566.344731	-1565.905123	-275.858303	16.56	9.51	16.56
G96LYP	-1701.69358	-1701.165081	-331.6380957	-6.26	-1570.409921	-1569.963136	-280.3619268	12.06	9.56	12.78
OLYP	-1701.831065	-1701.30303	-331.3472764	-5.97	-1570.562784	-1570.115806	-280.4828919	11.94	9.61	11.94
OVWN5	-1703.700825	-1703.157352	-341.034757	-15.66	-1573.101513	-1572.645109	-286.3980546	6.02	10.08	18.57
BP86	-1701.765494	-1701.223138	-340.3335778	-14.96	-1570.094379	-1569.620567	-297.3216476	-4.90	10.10	18.47
OPL	-1703.673112	-1703.128261	-341.8992958	-16.52	-1573.065063	-1572.607843	-286.910115	5.51	10.60	19.53
SPW91	-1695.571396	-1695.035466	-336.301245	-10.93	-1562.893055	-1562.411725	-302.0392648	-9.62	12.30	16.55
B2PLYP	-1700.883564	-1700.390645	-309.311097	16.06	-1569.37345	-1568.943036	-270.0891241	22.33	15.76	22.38
MPW2PLYP	-1700.933181	-1700.439517	-309.7788364	15.60	-1569.432499	-1569.000745	-270.9297296	21.49	15.82	22.04
B3P86	-1702.570065	-1702.02107	-344.4992061	-19.12	-1571.803689	-1571.315925	-306.0767557	-13.65	16.57	19.16
MPWB95	-1701.82684	-1701.288501	-337.8128278	-12.44	-1570.546365	-1570.017771	-331.6973451	-39.28	18.57	39.28
B2GPPLYP	-1700.678487	-1700.193933	-304.0622616	21.31	-1569.089054	-1568.667022	-264.8291566	27.59	20.99	29.82
DSD_BLYP	-1700.505599	-1700.025918	-301.0047525	24.37	-1568.851934	-1568.435187	-261.5125927	30.91	23.68	33.95
SVWN	-1698.349631	-1697.781279	-356.6467566	-31.27	-1566.753839	-1566.253702	-313.8402668	-21.42	25.48	37.74

Group C (Fe complexes with three water ligands and one amino acid sidechain ligand)

FUNCTIONAL	$[Fe(H_2O)_3(NH_2CH_3)]^{2+}$	$[Fe(H_2O)_3(NH_2CH_3)]^{3+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	$[Fe(H_2O)_3(OCH_3)]^{2+}$	$[Fe(H_2O)_3(OCH_3)]^{3+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)
CCSD(T)/CBS	-1587.158295	-1586.48742	-420.9802518	0.00	-1606.766919	-1606.332955	-272.3165482	0.00
BMK	-1587.742869	-1587.070941	-421.641258	-0.66	-1607.361874	-1606.927593	-272.5152519	-0.20
MPW1N	-1588.420565	-1587.747443	-422.3902785	-1.41	-1608.024034	-1607.589791	-272.4917516	-0.18
M062X	-1588.30009	-1587.627568	-422.014143	-1.03	-1607.916129	-1607.476703	-275.7439833	-3.43
MPW1K	-1588.414538	-1587.740743	-422.8132388	-1.83	-1608.016603	-1607.581501	-273.0304435	-0.71
MN12SX	-1588.069608	-1587.394208	-423.8195948	-2.84	-1607.694482	-1607.255011	-275.7722212	-3.46
MPWB1K	-1588.454448	-1587.782895	-421.4063937	-0.43	-1608.070696	-1607.638165	-271.4176811	0.90
MN12L	-1587.812133	-1587.142247	-420.3600904	0.62	-1607.441456	-1607.00665	-272.845209	-0.53
BB1K	-1588.450691	-1587.780211	-420.7327809	0.25	-1608.066804	-1607.635897	-270.3985177	1.92
M052X	-1588.360215	-1587.683876	-424.4090334	-3.43	-1607.979431	-1607.537788	-277.1351909	-4.82
B3LYP	-1588.498447	-1587.831919	-418.2530073	2.73	-1608.124694	-1607.696285	-268.83071	3.49
wB97X-D	-1588.385353	-1587.715755	-420.1792045	0.80	-1607.989151	-1607.567486	-264.5990112	7.72
M11L	-1588.455184	-1587.782979	-421.8154421	-0.84	-1608.081415	-1607.642822	-275.2209227	-2.90
MPW1B95	-1588.493001	-1587.826739	-418.0859894	2.89	-1608.118226	-1607.690949	-268.1203692	4.20
B3PW91	-1588.321412	-1587.656914	-416.9786608	4.00	-1607.934888	-1607.508904	-267.3091687	5.01
OPL	-1591.078819	-1590.411625	-418.6704895	2.31	-1610.712617	-1610.282835	-269.6924439	2.62
OVWN5	-1591.116672	-1590.450291	-418.1605375	2.82	-1610.751027	-1610.322395	-268.9705945	3.35
PBE1PBE	-1587.873394	-1587.210162	-416.1842399	4.80	-1607.480005	-1607.054641	-266.9200939	5.40
M11	-1588.238129	-1587.555997	-428.0443032	-7.06	-1607.863261	-1607.416156	-280.5623283	-8.25
B1LYP	-1588.330621	-1587.66691	-416.4854446	4.49	-1607.95817	-1607.533275	-266.6257291	5.69
MPW2PLYP	-1587.566372	-1586.901169	-417.4216387	3.56	-1607.169662	-1606.741135	-268.9051138	3.41
B97-2	-1588.588861	-1587.928638	-414.2960134	6.68	-1608.202802	-1607.782444	-263.7789252	8.54
B2PLYP	-1587.509028	-1586.84512	-416.6082921	4.37	-1607.111813	-1606.685196	-267.7060999	4.61

FUNCTIONAL	$[Fe(H_2O)_3(NH_2CH_3)]^{2+}$	$[Fe(H_2O)_3(NH_2CH_3)]^{3+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	$[Fe(H_2O)_3(OCH_3)]^{2+}$	$[Fe(H_2O)_3(OCH_3)]^{3+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)
B97-1	-1588.26462	-1587.603581	-414.8082494	6.17	-1607.874078	-1607.459191	-260.3457197	11.97
M06	-1588.217336	-1587.562569	-410.8725281	10.11	-1607.847349	-1607.427336	-263.5621394	8.75
TPSSTPSS	-1588.466632	-1587.810592	-411.6712976	9.31	-1608.090845	-1607.671626	-263.0644427	9.25
TPSS	-1588.466632	-1587.810592	-411.6712976	9.31	-1608.090845	-1607.671626	-263.0644427	9.25
B3P86	-1589.867557	-1589.182023	-430.1789339	-9.20	-1609.487341	-1609.040043	-280.6841969	-8.37
B2GPPLYP	-1587.233482	-1586.572252	-414.9280473	6.05	-1606.827694	-1606.401167	-267.6497244	4.67
SVWN	-1584.939023	-1584.282772	-411.8040284	9.18	-1604.511034	-1604.082831	-268.7018007	3.61
BHandH	-1584.550265	-1583.894696	-411.375979	9.60	-1604.118236	-1603.697997	-263.7042578	8.61
BP86	-1588.638357	-1587.98778	-408.2434323	12.74	-1608.275392	-1607.856144	-263.0819503	9.23
DSD_BLYP	-1587.002795	-1586.344848	-412.8684478	8.11	-1606.59255	-1606.167973	-266.4258297	5.89
HCTH407	-1588.959402	-1588.308422	-408.4959672	12.48	-1608.575639	-1608.160345	-260.6010095	11.72
N12	-1588.718661	-1588.068004	-408.2939154	12.69	-1608.348173	-1607.934137	-259.8119789	12.50
M05	-1588.280757	-1587.629822	-408.4678673	12.51	-1607.912258	-1607.499423	-259.0578316	13.26
BPW91	-1588.540671	-1587.893031	-406.4002232	14.58	-1608.173681	-1607.758514	-260.521322	11.80
VSXC	-1589.057424	-1588.411887	-405.0804011	15.90	-1608.696578	-1608.278162	-262.5599815	9.76
BPBE	-1588.331983	-1587.685562	-405.6351321	15.35	-1607.961738	-1607.547712	-259.8054842	12.51
MPWB95	-1588.607077	-1587.962817	-404.2790964	16.70	-1608.258569	-1607.844449	-259.8638488	12.45
G96LYP	-1588.511658	-1587.867375	-404.2931903	16.69	-1608.156728	-1607.745662	-257.947353	14.37
B97D3	-1588.767978	-1588.124563	-403.7488132	17.23	-1608.386662	-1607.976686	-257.2640201	15.05
OTPSS-D	-1588.549993	-1587.907324	-403.2805655	17.70	-1608.160389	-1607.753632	-255.2437467	17.07
M06L	-1588.359022	-1587.718607	-401.8667989	19.11	-1607.984397	-1607.57801	-255.0116434	17.30
OLYP	-1588.622257	-1587.980403	-402.77003	18.21	-1608.250474	-1607.843718	-255.243276	17.07
SPW91	-1581.14662	-1580.531538	-385.9700041	35.01	-1600.693893	-1600.304675	-244.2378244	28.08
B1B95	-1588.490582	-1587.826428	-416.7634941	4.22	-1608.116236	-1607.691326	-266.6346021	5.68

FUNCTIONAL	$[Fe(H_2O)_3(COOH)]^+$	$[Fe(H_2O)_3(COOH)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	$[Fe(H_2O)_3(SCH_3)]^+$	$[Fe(H_2O)_3(SCH_3)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	MUE (kcal/mol)	MaxE (kcal/mol)
CCSD(T)/CBS	-1680.782248	-1680.298346	-303.6529423	0.00	-1929.374019	-1928.95426	-263.4024515	0.00	0.00	0.00
BMK	-1681.424969	-1680.944497	-301.5006545	2.15	-1930.256227	-1929.838919	-261.8643118	1.54	1.14	2.15
MPW1N	-1682.077483	-1681.595374	-302.5278563	1.13	-1931.03974	-1930.626887	-259.0688256	4.33	1.76	4.33
M062X	-1681.98559	-1681.502044	-303.4302402	0.22	-1930.891659	-1930.47638	-260.5916282	2.81	1.87	3.43
MPW1K	-1682.068703	-1681.585686	-303.09825	0.55	-1931.033397	-1930.620967	-258.8033953	4.60	1.93	4.60
MN12SX	-1681.736768	-1681.252426	-303.9294618	-0.28	-1930.67663	-1930.254391	-264.9586181	-1.56	2.03	3.46
MPWB1K	-1682.133823	-1681.654353	-300.8724734	2.78	-1931.121124	-1930.70817	-259.1326182	4.27	2.09	4.27
MN12L	-1681.480598	-1681.002582	-299.9595097	3.69	-1930.438694	-1930.01223	-267.6098336	-4.21	2.26	4.21
BB1K	-1682.130091	-1681.65221	-299.8747269	3.78	-1931.112059	-1930.699672	-258.7767638	4.63	2.64	4.63
M052X	-1682.062957	-1681.575545	-305.8551695	-2.20	-1930.94651	-1930.53161	-260.354166	3.05	3.37	4.82
B3LYP	-1682.204319	-1681.731686	-296.5822468	7.07	-1931.106128	-1930.687851	-262.4727514	0.93	3.55	7.07
wB97X-D	-1682.05579	-1681.579465	-298.8984663	4.75	-1930.986827	-1930.568848	-262.2863057	1.12	3.60	7.72
M11L	-1682.135151	-1681.65705	-300.0129045	3.64	-1931.070292	-1930.637221	-271.7564484	-8.35	3.93	8.35
MPW1B95	-1682.189353	-1681.716411	-296.7756452	6.88	-1931.16228	-1930.746895	-260.6584141	2.74	4.18	6.88
B3PW91	-1681.984512	-1681.513493	-295.5691388	8.08	-1930.899306	-1930.48141	-262.2335761	1.17	4.57	8.08
OPL	-1684.985688	-1684.517575	-293.7458344	9.91	-1934.033408	-1933.606649	-267.7950745	-4.39	4.81	9.91
OVWN5	-1685.028789	-1684.561468	-293.2484954	10.40	-1934.074333	-1933.648669	-267.108648	-3.71	5.07	10.40
PBE1PBE	-1681.492759	-1681.021713	-295.5859812	8.07	-1930.388788	-1929.973595	-260.5376498	2.86	5.28	8.07
M11	-1681.934676	-1681.441088	-309.7316163	-6.08	-1930.838079	-1930.42245	-260.8111436	2.59	5.99	8.25
B1LYP	-1682.027428	-1681.557181	-295.0846324	8.57	-1930.942656	-1930.531702	-257.8775676	5.52	6.07	8.57
MPW2PLYP	-1681.15823	-1680.682916	-298.2642047	5.39	-1930.083393	-1929.686581	-249.0034084	14.40	6.69	14.40
B97-2	-1682.258632	-1681.793205	-292.0596659	11.59	-1931.225673	-1930.811106	-260.144446	3.26	7.52	11.59
B2PLYP	-1681.095646	-1680.622098	-297.1562863	6.50	-1930.013892	-1929.618234	-248.2789172	15.12	7.65	15.12

FUNCTIONAL	$[Fe(H_2O)_3(COOH)]^+$	$[Fe(H_2O)_3(COOH)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	$[Fe(H_2O)_3(SCH_3)]^+$	$[Fe(H_2O)_3(SCH_3)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	MUE (kcal/mol)	MaxE (kcal/mol)
B97-1	-1681.94814	-1681.480545	-293.4202824	10.23	-1930.834565	-1930.418674	-260.9757143	2.43	7.70	11.97
M06	-1681.904379	-1681.4429	-289.5824148	14.07	-1930.831982	-1930.412917	-262.9674172	0.44	8.34	14.07
TPSSTPSS	-1682.168887	-1681.706912	-289.8939796	13.76	-1931.076086	-1930.660021	-261.0847127	2.32	8.66	13.76
TPSS	-1682.168887	-1681.706912	-289.8939796	13.76	-1931.076086	-1930.660021	-261.0847127	2.32	8.66	13.76
B3P86	-1683.668871	-1683.176596	-308.9068554	-5.25	-1932.634778	-1932.195288	-275.7841502	-12.38	8.80	12.38
B2GPPLYP	-1680.781647	-1680.309188	-296.4724954	7.18	-1929.704191	-1929.313554	-245.1287814	18.27	9.04	18.27
SVWN	-1678.229152	-1677.765496	-290.9482209	12.70	-1926.813486	-1926.375141	-275.0657207	-11.66	9.29	12.70
BHandH	-1677.821651	-1677.354291	-293.272962	10.38	-1926.538971	-1926.134104	-254.0581928	9.34	9.49	10.38
BP86	-1682.358373	-1681.902574	-286.0184176	17.63	-1931.285809	-1930.865723	-263.607546	-0.21	9.95	17.63
DSD_BLYP	-1680.523305	-1680.053431	-294.850496	8.80	-1929.443039	-1929.055984	-242.8806341	20.52	10.83	20.52
HCTH407	-1682.619243	-1682.165434	-284.7692719	18.88	-1931.619346	-1931.202878	-261.3377497	2.06	11.29	18.88
N12	-1682.414745	-1681.962546	-283.7588559	19.89	-1931.350763	-1930.933812	-261.6409184	1.76	11.71	19.89
M05	-1681.9748	-1681.519081	-285.9684051	17.68	-1930.923486	-1930.509984	-259.4767507	3.93	11.85	17.68
BPW91	-1682.250611	-1681.798331	-283.8099101	19.84	-1931.164707	-1930.749156	-260.7623798	2.64	12.21	19.84
VSXC	-1682.791715	-1682.340412	-283.1971595	20.46	-1931.795056	-1931.380379	-260.2133717	3.19	12.33	20.46
BPBE	-1682.018714	-1681.567536	-283.1185953	20.53	-1930.920049	-1930.505245	-260.2937557	3.11	12.87	20.53
MPWB95	-1682.349427	-1681.90091	-281.4488299	22.20	-1931.289652	-1930.874699	-260.3870726	3.02	13.59	22.20
G96LYP	-1682.237031	-1681.789251	-280.9862612	22.67	-1931.151438	-1930.741814	-257.042572	6.36	15.02	22.67
B97D3	-1682.422612	-1681.974786	-281.0149824	22.64	-1931.414657	-1931.003888	-257.7614721	5.64	15.14	22.64
OTPSS-D	-1682.202004	-1681.756168	-279.7666461	23.89	-1931.152703	-1930.741693	-257.9128148	5.49	16.04	23.89
M06L	-1682.059385	-1681.613787	-279.6170227	24.04	-1930.95868	-1930.5483	-257.5169378	5.89	16.58	24.04
OLYP	-1682.317015	-1681.872256	-279.0905673	24.56	-1931.254534	-1930.84688	-255.8069742	7.60	16.86	24.56
SPW91	-1674.069844	-1673.644754	-266.748583	36.90	-1922.53054	-1922.124203	-254.9801173	8.42	27.10	36.90
B1B95	-1682.18786	-1681.717692	-295.0349587	8.62	-	-	-	-	-	-

Group D (Fe complexes with three water ligands and one amino acid sidechain ligand)

FUNCTIONAL	$[Fe(H_2O)_5(NH_2CH_3)]^{2+}$	$[Fe(H_2O)_5(NH_2CH_3)]^{3+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	$[Fe(H_2O)_5(COOH)]^{+}$	$Fe(H_2O)_5(COOH)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)
CCSD(T)/CBS	-1740.000832	-1739.404676	-374.0937041	0	-1833.600745	-1833.171583	-269.3032719	0
wB97X-D	-1741.341225	-1740.741782	-376.1561995	-2.06	-1834.994106	-1834.561547	-271.4352325	-2.13
MPWB1K	-1741.378168	-1740.781141	-374.640312	-0.55	-1835.042776	-1834.612112	-270.2462462	-0.94
B3PW91	-1741.254505	-1740.658576	-373.9507731	0.14	-1834.90054	-1834.471756	-269.0657061	0.24
BB1K	-1741.374462	-1740.777802	-374.410085	-0.32	-1835.039057	-1834.609096	-269.8046049	-0.50
MPW1B95	-1741.425876	-1740.830535	-373.5821426	0.51	-1835.106818	-1834.678478	-268.7876566	0.52
M062X	-1741.244127	-1740.647936	-374.1154818	-0.02	-1834.914557	-1834.482536	-271.0973186	-1.79
B3LYP	-1741.494281	-1740.895168	-375.9486256	-1.85	-1835.182697	-1834.751046	-270.8654413	-1.56
BMK	-1740.66183	-1740.062876	-375.8491716	-1.76	-1834.328776	-1833.895357	-271.9744202	-2.67
MN12L	-1740.657468	-1740.054985	-378.0644625	-3.97	-1834.308848	-1833.875143	-272.1539193	-2.85
MPW1N	-1741.360897	-1740.761959	-375.8392256	-1.75	-1835.001847	-1834.568802	-271.7399136	-2.44
MPW1K	-1741.353421	-1740.754394	-375.8956324	-1.80	-1834.991717	-1834.558478	-271.8612739	-2.56
PBE1PBE	-1740.702431	-1740.109425	-372.1166629	1.98	-1834.305108	-1833.878346	-267.7971452	1.51
OVWN5	-1744.678391	-1744.069309	-382.2048273	-8.11	-1838.568696	-1838.134698	-272.3378863	-3.03
B1LYP	-1741.264121	-1740.668719	-373.6199187	0.47	-1834.943746	-1834.515677	-268.6170681	0.69
OPL	-1744.631665	-1744.021651	-382.7900427	-8.70	-1838.516674	-1838.081731	-272.930864	-3.63
M052X	-1741.343124	-1740.743857	-376.0462159	-1.95	-1835.030797	-1834.595227	-273.3243565	-4.02
B97-1	-1741.21098	-1740.617561	-372.3760503	1.72	-1834.8726	-1834.450957	-264.5848985	4.72
MN12SX	-1740.95563	-1740.349898	-380.1026388	-6.01	-1834.605503	-1834.16759	-274.7945362	-5.49
B97-2	-1741.525844	-1740.932972	-372.0330033	2.06	-1835.177826	-1834.753506	-266.2649488	3.04
M11L	-1741.410025	-1740.802638	-381.1413554	-7.05	-1835.071048	-1834.63454	-273.9125589	-4.61
HCTH407	-1741.925547	-1741.335125	-370.4955672	3.60	-1835.56451	-1835.145928	-262.6643991	6.64
MPW2PLYP	-1740.305012	-1739.714763	-370.3868574	3.71	-1833.880882	-1833.456628	-266.2236335	3.08

FUNCTIONAL	$[Fe(H_2O)_5(NH_2CH_3)]^{2+}$	$[Fe(H_2O)_5(NH_2CH_3)]^{3+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	$[Fe(H_2O)_5(COOH)]^{+}$	$Fe(H_2O)_5(COOH)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)
BP86	-1741.628191	-1741.042738	-367.37729	6.72	-1835.329964	-1834.911419	-262.6407043	6.66
TPSS	-1741.446722	-1740.859152	-368.7053576	5.39	-1835.131745	-1834.711303	-263.831272	5.47
B2PLYP	-1740.233926	-1739.644248	-370.0286436	4.07	-1833.804491	-1833.381115	-265.6725421	3.63
TPSSTPSS	-1741.446722	-1740.859152	-368.7053576	5.39	-1835.131745	-1834.711303	-263.831272	5.47
M11	-1741.199776	-1740.591319	-381.8125459	-7.72	-1834.881365	-1834.436331	-279.2633332	-9.96
VSXC	-1742.144576	-1741.562576	-365.2105307	8.88	-1835.860932	-1835.449691	-258.058008	11.25
N12	-1741.618786	-1741.034388	-366.7158196	7.38	-1835.295691	-1834.881425	-259.9556347	9.35
BPW91	-1741.508379	-1740.925343	-365.8610637	8.23	-1835.199774	-1834.784635	-260.504122	8.80
MPWB95	-1741.564088	-1740.98437	-363.7783342	10.32	-1835.288788	-1834.876896	-258.4660587	10.84
BPBE	-1741.251071	-1740.669339	-365.0425715	9.05	-1834.91929	-1834.505345	-259.75475	9.55
BHandH	-1736.618854	-1736.04456	-360.3750923	13.72	-1829.876148	-1829.463347	-259.0368414	10.27
B97D3	-1741.688058	-1741.108968	-363.3849923	10.71	-1835.32299	-1834.912011	-257.8929604	11.41
G96LYP	-1741.455959	-1740.875454	-364.2721842	9.82	-1835.162606	-1834.750684	-258.4847835	10.82
OTPSS-D	-1741.450278	-1740.86913	-364.6759804	9.42	-1835.082248	-1834.672712	-256.9874013	12.32
M05	-1741.203873	-1740.618455	-367.3555154	6.74	-1834.880461	-1834.463467	-261.6678448	7.64
DSD_BLYP	-1739.598415	-1739.017464	-364.552361	9.54	-1833.103274	-1832.688291	-260.4060485	8.90
M06	-1741.153121	-1740.567352	-367.5754011	6.52	-1834.822595	-1834.404629	-262.2777779	7.03
B3P86	-1743.217893	-1742.601718	-386.6558847	-12.56	-1837.002428	-1836.552747	-282.1786613	-12.88
OLYP	-1741.549836	-1740.968822	-364.5916933	9.50	-1835.224142	-1834.814787	-256.8745939	12.43
B2GPPLYP	-1739.888006	-1739.303231	-366.9516126	7.14	-1833.420452	-1833.00166	-262.7960004	6.51
M06L	-1741.336223	-1740.763602	-359.3251747	14.77	-1835.018694	-1834.615868	-252.7771696	16.53
SVWN	-1737.222444	-1736.638482	-366.4420058	7.65	-1830.491999	-1830.074269	-262.1295664	7.17
SPW91	-1732.426472	-1731.888119	-337.8220396	36.27	-1825.331235	-1824.955116	-236.0181201	33.29
B1B95	-1741.423928	-1740.829705	-372.8805555	1.21	-1835.105694	-1834.678913	-267.8092813	1.49

FUNCTIONAL	$[Fe(H_2O)_5(OCH_3)]^+$	$[Fe(H_2O)_5(OCH_3)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	$[Fe(H_2O)_5(SCH_3)]^+$	$[Fe(H_2O)_5(SCH_3)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	MUE (kcal/mol)	MaxE (kcal/mol)
CCSD(T)/CBS	-1759.584479	-1759.196892	-243.2144608	0	-2082.184746	-2081.802672	-239.7554738	0	0.00	0.00
wB97X-D	-1760.925237	-1760.538773	-242.5101148	0.70	-2083.91246	-2083.529764	-240.1454955	-0.39	1.32	2.13
MPWB1K	-1760.96858	-1760.581047	-243.1810794	0.03	-2084.012982	-2083.637858	-235.3943128	4.36	1.47	4.36
B3PW91	-1760.841085	-1760.45851	-240.0693534	3.15	-2083.798041	-2083.420715	-236.775612	2.98	1.63	3.15
BB1K	-1760.964732	-1760.57851	-242.3582951	0.86	-2084.003805	-2083.629454	-234.908407	4.85	1.63	4.85
MPW1B95	-1761.02451	-1760.641401	-240.4042867	2.81	-2084.062186	-2083.685901	-236.122682	3.63	1.87	3.63
M062X	-1760.834103	-1760.440643	-246.8998017	-3.69	-2083.803825	-2083.426228	-236.9456232	2.81	2.08	3.69
B3LYP	-1761.092711	-1760.707503	-241.7215674	1.49	-2084.066172	-2083.689601	-236.3017168	3.45	2.09	3.45
BMK	-1760.254191	-1759.864633	-244.4512718	-1.24	-2083.140556	-2082.763545	-236.5783732	3.18	2.21	3.18
MN12L	-1760.255888	-1759.867178	-243.9196959	-0.71	-2083.244892	-2082.864912	-238.4410162	1.31	2.21	3.97
MPW1N	-1760.938758	-1760.548921	-244.6264662	-1.41	-2083.948059	-2083.572993	-235.3573023	4.40	2.50	4.40
MPW1K	-1760.929991	-1760.539487	-245.045241	-1.83	-2083.94043	-2083.565619	-235.1974944	4.56	2.69	4.56
PBE1PBE	-1760.282674	-1759.900812	-239.6217571	3.59	-2083.184646	-2082.808721	-235.896584	3.86	2.73	3.86
OVWN5	-1764.281372	-1763.894417	-242.8184479	0.40	-2087.598095	-2087.215207	-240.2654063	-0.51	3.01	8.11
B1LYP	-1760.864085	-1760.48225	-239.6045383	3.61	-2083.840868	-2083.470915	-232.1495109	7.61	3.09	7.61
OPL	-1764.234036	-1763.84596	-243.5214093	-0.31	-2087.548243	-2087.164296	-240.9305664	-1.18	3.45	8.70
M052X	-1760.93589	-1760.540457	-248.1381103	-4.92	-2083.897604	-2083.520511	-236.6300235	3.13	3.51	4.92
B97-1	-1760.801352	-1760.420879	-238.75014	4.46	-2083.7461	-2083.371159	-235.2791648	4.48	3.84	4.72
MN12SX	-1760.550394	-1760.155985	-247.4955217	-4.28	-2083.526417	-2083.145518	-239.0177729	0.74	4.13	6.01
B97-2	-1761.111907	-1760.734408	-236.8844347	6.33	-2084.127577	-2083.754187	-234.3063429	5.45	4.22	6.33
M11L	-1761.005409	-1760.61177	-247.0119942	-3.80	-2083.986199	-2083.594709	-245.663476	-5.91	5.34	7.05
HCTH407	-1761.511634	-1761.138294	-234.2745282	8.94	-2084.550987	-2084.17438	-236.3249158	3.43	5.65	8.94
MPW2PLYP	-1759.881935	-1759.49815	-240.8284832	2.39	-2082.789192	-2082.430575	-225.03556	14.72	5.97	14.72

FUNCTIONAL	$[Fe(H_2O)_5(OCH_3)]^+$	$[Fe(H_2O)_5(OCH_3)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	$[Fe(H_2O)_5(SCH_3)]^+$	$[Fe(H_2O)_5(SCH_3)]^{2+}$	ΔEel (kcal/mol)	ΔEel - CCSD(T)/CBS (kcal/mol)	MUE (kcal/mol)	MaxE (kcal/mol)
BP86	-1761.237762	-1760.862045	-235.766062	7.45	-2084.239429	-2083.862267	-236.6728699	3.08	5.98	7.45
TPSS	-1761.009839	-1760.63652	-234.2611434	8.95	-2084.021761	-2083.647394	-234.9190621	4.84	6.16	8.95
B2PLYP	-1759.810322	-1759.428045	-239.882594	3.33	-2082.705853	-2082.348484	-224.2527167	15.50	6.63	15.50
TPSSTPSS	-1761.043912	-1760.63652	-255.6428302	-12.43	-2084.021761	-2083.647394	-234.9190621	4.84	7.03	12.43
M11	-1760.7971	-1760.395728	-251.8649589	-8.65	-2083.767429	-2083.390124	-236.7622837	2.99	7.33	9.96
VSXC	-1761.746143	-1761.375332	-232.6875564	10.53	-2084.854324	-2084.471891	-239.9803977	-0.22	7.72	11.25
N12	-1761.217982	-1760.84996	-230.9373066	12.28	-2084.216739	-2083.839507	-236.7169775	3.04	8.01	12.28
BPW91	-1761.113519	-1760.741946	-233.1654614	10.05	-2084.096203	-2083.723519	-233.8626685	5.89	8.24	10.05
MPWB95	-1761.18744	-1760.817395	-232.2066456	11.01	-2084.214674	-2083.838072	-236.3215335	3.43	8.90	11.01
BPBE	-1760.853078	-1760.482647	-232.4487074	10.77	-2083.80313	-2083.431081	-233.464225	6.29	8.91	10.77
BHandH	-1756.163182	-1755.789792	-234.305797	8.91	-2078.58096	-2078.209572	-233.0495165	6.71	9.90	13.72
B97D3	-1761.276925	-1760.909639	-230.4757796	12.74	-2084.302545	-2083.929858	-233.8644632	5.89	10.19	12.74
G96LYP	-1761.072706	-1760.704996	-230.7415111	12.47	-2084.058131	-2083.692512	-229.4297211	10.33	10.86	12.47
OTPSS-D	-1761.032025	-1760.66787	-228.5107898	14.70	-2084.017832	-2083.648432	-231.8019396	7.95	11.10	14.70
M05	-1760.784993	-1760.435492	-219.3150445	23.90	-2083.809856	-2083.437749	-233.5005955	6.25	11.13	23.90
DSD_BLYP	-1759.162058	-1758.783912	-237.2902264	5.92	-2082.006331	-2081.656935	-219.2491622	20.51	11.22	20.51
M06	-1760.719943	-1760.378267	-214.4053215	28.81	-2083.730887	-2083.353563	-236.7743507	2.98	11.33	28.81
B3P86	-1762.811027	-1762.407348	-253.3122847	-10.10	-2085.950543	-2085.551889	-250.158892	-10.40	11.48	12.88
OLYP	-1761.148469	-1760.784359	-228.4824389	14.73	-2084.145524	-2083.78043	-229.1004543	10.66	11.83	14.73
B2GPPLYP	-1759.428955	-1759.075373	-221.8759739	21.34	-2082.326253	-2081.973436	-221.3960545	18.36	13.34	21.34
M06L	-1760.931518	-1760.571099	-226.1662319	17.05	-2083.897958	-2083.530842	-230.3685131	9.39	14.43	17.05
SVWN	-1756.699253	-1756.387352	-195.7209659	47.49	-2079.062353	-2078.666881	-248.1626334	-8.41	17.68	47.49
SPW91	-1751.888966	-1751.609133	-175.5984775	67.62	-2073.780456	-2073.414104	-229.8896668	9.87	36.76	67.62
B1B95	-1761.022913	-1760.641886	-239.0982572	4.12	-	-	-	-	-	-

II. Impact of Polarity of Fullerene Derivatives on the Morphology of Bulk Heterojunction Solar Cells from Solvent Evaporation Simulations

Simulation Parameter

```
; STANDARD MD INPUT OPTIONS FOR MARTINI 2.x
; Updated 15 Jul 2015 by DdJ
;
; for use with GROMACS 5
; For a thorough comparison of different mdp options in combination with the
Martini force field, see:
; D.H. de Jong et al., Martini straight: boosting performance using a shorter
cutoff and GPUs, submitted.

title                    = Martini ('new' parameters) NPT run (semiiso)

; TIMESTEP IN MARTINI
; Most simulations are numerically stable with dt=40 fs,
; however better energy conservation is achieved using a
; 20-30 fs timestep.
; Time steps smaller than 20 fs are not required unless specifically stated in
the itp file.

integrator               = md
dt                       = 0.02
nsteps                   = 1500000 ; 30 ns
nstcomm                  = 100
comm-grps                      =

nstxout                  = 10000000
nstvout                  = 0
nstfout                  = 0
nstlog                   = 1000
nstenergy                = 100
nstxout-compressed       = 1000
compressed-x-precision   = 100
compressed-x-grps        = System
energygrps               = System

; NEIGHBOURLIST and MARTINI
; To achieve faster simulations in combination with the Verlet-neighborlist
; scheme, Martini can be simulated with a straight cutoff. In order to
; do so, the cutoff distance is reduced 1.1 nm.
; Neighborlist length should be optimized depending on your hardware setup:
; updating ever 20 steps should be fine for classic systems, while updating
; every 30-40 steps might be better for GPU based systems.
; The Verlet neighborlist scheme will automatically choose a proper
neighborlist
; length, based on a energy drift tolerance.
;
; Coulomb interactions can alternatively be treated using a reaction-field,
; giving slightly better properties.
; Please realize that electrostVatic interactions in the Martini model are
; not considered to be very accurate to begin with, especially as the
; screening in the system is set to be uniform across the system with
; a screening constant of 15. When using PME, please make sure your
; system properties are still reasonable.
;
; With the polarizable water model, the relative electrostatic screening
; (epsilon_r) should have a value of 2.5, representative of a low-dielectric
; apolar solvent. The polarizable water itself will perform the explicit
screening
; in aqueous environment. In this case, the use of PME is more realistic.


cutoff-scheme            = Verlet
nstlist                  = 20
```

```
ns_type                  = grid
pbc                      = xyz
verlet-buffer-tolerance  = 0.005

coulombtype              = cutoff
coulomb-modifier         = Potential-shift-verlet
rcoulomb                 = 1.1
epsilon_r                = 15   ; 2.5 (with polarizable water)
vdw_type                 = cutoff
vdw-modifier             = Potential-shift-verlet
rvdw                     = 1.1

; MARTINI and TEMPERATURE/PRESSURE
; normal temperature and pressure coupling schemes can be used.
; It is recommended to couple individual groups in your system separately.
; Good temperature control can be achieved with the velocity rescale (V-
rescale)
; thermostat using a coupling constant of the order of 1 ps. Even better
; temperature control can be achieved by reducing the temperature coupling
; constant to 0.1 ps, although with such tight coupling (approaching
; the time step) one can no longer speak of a weak-coupling scheme.
; We therefore recommend a coupling time constant of at least 0.5 ps.
; The Berendsen thermostat is less suited since it does not give
; a well described thermodynamic ensemble.
;
; Pressure can be controlled with the Parrinello-Rahman barostat,
; with a coupling constant in the range 4-8 ps and typical compressibility
; in the order of 10e-4 - 10e-5 bar-1. Note that, for equilibration purposes,
; the Berendsen barostat probably gives better results, as the Parrinello-
; Rahman is prone to oscillating behaviour. For bilayer systems the pressure
; coupling should be done semiisotropic.

tcoupl                   = v-rescale
tc-grps                  = System
tau_t                    = 1.0
ref_t                    = 298

Pcoupl                   = parrinello-rahman
Pcoupltype               = semiisotropic
tau_p                    = 15.0 15.0  ;parrinello-rahman is more stable with
larger tau-p, DdJ, 20130422
compressibility          =  0  3e-4
ref_p                    =  1     1

gen_vel                  = no
gen_temp                 = 298
gen_seed                 = 473529

; MARTINI and CONSTRAINTS
; for ring systems and stiff bonds constraints are defined
; which are best handled using Lincs.

constraints              = none
constraint_algorithm     = Lincs
```